

Covariance Eigenvector Sparsity for Compression and Denoising

*Ioannis D. Schizas and Georgios B. Giannakis**

Abstract

Sparsity in the eigenvectors of signal covariance matrices is exploited in this paper for compression and denoising. Dimensionality reduction (DR) and quantization modules present in many practical compression schemes such as transform codecs, are designed to capitalize on this form of sparsity and achieve improved reconstruction performance compared to existing sparsity-agnostic codecs. Using training data that may be noisy a novel sparsity-aware linear DR scheme is developed to fully exploit sparsity in the covariance eigenvectors and form noise-resilient estimates of the principal covariance eigenbasis. Sparsity is effected via norm-one regularization, and the associated minimization problems are solved using computationally efficient coordinate descent iterations. The resulting eigenspace estimator is shown capable of identifying a subset of the unknown support of the eigenspace basis vectors even when the observation noise covariance matrix is unknown, as long as the noise power is sufficiently low. It is proved that the sparsity-aware estimator is asymptotically normal, and the probability to correctly identify the signal subspace basis support approaches one, as the number of training data grows large. Simulations using synthetic data and images, corroborate that the proposed algorithms achieve improved reconstruction quality relative to alternatives.

Index Terms

PCA, data compression, subspace estimation, denoising, quantization.

† Manuscript received October 7, 2010; accepted January 7, 2012. First published XXXXXX XX, 20XX; current version published XXXXXX XX, 20XX. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Maja Bystrom. Work in this paper was supported by the NSF Grant CCF-1016605. Part of the paper was presented in the *3rd Intl. Workshop on Comp. Advances in Multi-Sensor Adap. Proc.*, Aruba, Dutch Antilles, Dec. 2009.

* Ioannis Schizas is with the Dept. of Electrical Engineering, University of Texas at Arlington, 416 Yates Street, Arlington, TX 76011. Tel/fax: (817) 272-3467/272-2253. Georgios Giannakis is with the Dept. of Electrica and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455; Emails: {schizas@uta.edu, georgios@umn.edu}.

I. INTRODUCTION

Data compression has well-appreciated impact in audio, image and video processing since the increasing data rates cannot be matched by the computational and storage capabilities of existing processors. The cornerstone modules of compression are those performing dimensionality reduction (DR) and quantization, as in e.g., transform codecs [16]. DR projects the data onto a space of lower dimension while minimizing an appropriate figure of merit quantifying information loss. Quantization amounts to digitizing the analog-amplitude DR data. Typically, DR relies on training vectors to find parsimonious data representations with reduced redundancy without inducing, e.g., mean-square error (MSE) distortion in the reconstruction process. One such property that promotes parsimony is *sparsity*.

Sparsity is an attribute characterizing many natural and man-made signals [2], and has been successfully exploited in statistical inference tasks using the least-absolute shrinkage and selection operator (Lasso) [29], [35]. In parallel, recent results in compressive sampling rely on sparsity to solve under-determined systems of linear equations, as well as sample continuous signals at sub-Nyquist rates [7]. These tasks take advantage of sparsity present in deterministic signal descriptors. But what if sparsity is present in statistical descriptors such as the signal covariance matrix? The latter is instrumental in compression when the distortion metric is MSE. In bio-informatics and imaging applications the data input to the DR module have covariance matrices whose eigenvectors admit a sparse representation over a certain domain, such as the wavelet domain [18].

The ‘workhorse’ for DR is principal component analysis (PCA) [5], [19], which relies on the covariance matrix to project the data on the subspace spanned by its principal eigenvectors. So far, sparsity has been exploited for interpreting each principal component, but not for reconstructing reduced-dimension signal renditions. Specifically, the standard PCA criterion has been regularized with an ℓ_1 -norm penalty term to induce sparsity as in the Lasso, and perform variable selection of the data entries that significantly contribute to each principal component [20]. However, the nonconvex formulation of [20] does not lend itself to efficient optimization. The sparse PCA formulation in [36] leads to a cost minimized using the block coordinate descent optimization method [3]; see also [23] which includes a weighted ℓ_1 -norm sparsity-imposing penalty. PCA with cardinality constraints on the number of nonzero elements per principal eigenvector has also been considered using relaxation techniques [8], [9]. Alternative approaches augment the standard singular value decomposition (SVD) cost, or the maximum likelihood criterion with ℓ_1 (or ℓ_0) penalties to effect sparsity in the principal eigenvectors [18], [27], [31]–[33]. Sparsity has also been exploited to render PCA robust to outliers [6], as well as to reduce complexity at the encoder [12]. In all the aforementioned schemes sparsity is not exploited for reconstructing signals that have been compressed

by DR and quantization.

When dealing with noisy data, pertinent reconstruction techniques have been developed to perform joint denoising and signal reconstruction [25], [26], [28]. However, existing approaches rely on the noise second-order statistics being available. Here, the a priori knowledge that the signal covariance eigenvectors are sparse is exploited, and joint denoising-reconstruction schemes are introduced without requiring availability of the noise covariance.

The standard PCA cost is augmented with pertinent ℓ_1 - and ℓ_2 -norm regularization terms, that fully exploit the sparsity present in the covariance eigenvectors when performing not only feature extraction (as in [9], [20], [33], [36]) but also reconstruction. The resulting bilinear cost is minimized via coordinate descent which leads to an efficient sparse (S-) PCA algorithm for DR and reconstruction. Its large-sample performance is analyzed in the presence of observation noise. If the ensemble covariance matrix of the noisy training data is known, then the sparsity-aware estimates do better in terms of identifying the support of the principal basis vectors when compared to the standard sparsity-agnostic PCA. As the number of training data used to design the DR module grows large, the novel sparsity-aware signal covariance eigenspace estimators: i) are asymptotically normal; and ii) identify the true principal eigenvectors' support (indices of nonzero elements) with probability one. The last two properties, known as 'oracle' properties [35], combined with the noise-resilience enable the novel S-PCA to attain an improved trade-off between reconstruction performance and reduced-dimensionality-a performance advantage also corroborated via numerical examples. The proposed sparsity-aware DR scheme is finally combined with a vector quantizer (VQ) to obtain a sparsity-aware transform coder (SATC), which improves reconstruction performance when compared to standard TC schemes that rely on the discrete cosine transform (DCT) or PCA transform. The merits of SATC are also demonstrated in the compression and reconstruction/denoising of noisy images that have been extracted from [1].

The rest of the paper is organized as follows. After stating the problem setting in Sec. II, the proposed sparsity-aware PCA formulation is introduced in Sec. III-A. Coordinate descent is employed in Sec. III-B to minimize the associated bilinear cost, while a computationally simpler element-wise algorithm is derived in Sec. III-C. A cross-validation scheme is outlined in Sec. III-D for selecting the sparsity-controlling coefficients that weigh the ℓ_1 -norm based regularization terms. Asymptotic properties are derived both in the noiseless and noisy cases, establishing the potential of the novel estimators to recover the underlying signal covariance principal eigenvectors (Secs. IV-A, IV-B). S-PCA is combined with vector quantization in Sec. V. Synthetic numerical examples (Sec. VI-A) illustrate the performance advantage of SATC, while tests using real noisy images corroborate the potential of SATC in practical settings (Sec. VI-B).

II. PRELIMINARIES AND PROBLEM FORMULATION

Consider a collection of n training data vectors $\{\mathbf{x}_t = \mathbf{s}_t + \mathbf{w}_t\}_{t=1}^n$, each containing the signal of interest $\mathbf{s}_t \in \mathbb{R}^{p \times 1}$, in additive zero-mean possibly colored noise \mathbf{w}_t , assumed independent of \mathbf{s}_t . It is also assumed that \mathbf{s}_t lies in a linear subspace of reduced dimension, say $r \leq p$, spanned by an unknown orthogonal basis $\{\mathbf{u}_{s,\rho}\}_{\rho=1}^r$. Many images and audio signals lie on such a low-dimensional subspace. Accordingly, \mathbf{s}_t for such signals can be expressed as

$$\mathbf{s}_t = \boldsymbol{\mu}_s + \sum_{\rho=1}^r \pi_{t,\rho} \mathbf{u}_{s,\rho}, \quad t = 1, \dots, n \quad (1)$$

where $\boldsymbol{\mu}_s$ denotes the mean of \mathbf{s}_t , and $\pi_{t,\rho}$ are zero-mean independent projection coefficients.

The covariance matrix of the noisy \mathbf{x}_t is given by $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_s + \boldsymbol{\Sigma}_w$, where $\boldsymbol{\Sigma}_s$ ($\boldsymbol{\Sigma}_w$) denotes the signal (noise) covariance matrix. Consider the eigen-decomposition $\boldsymbol{\Sigma}_w = \mathbf{U}_w \mathbf{D}_w \mathbf{U}_w^T$, where \mathbf{U}_w denotes the eigen-matrix containing the eigenspace basis, and $\mathbf{D}_w := \text{diag}(d_{w,1}, \dots, d_{w,p})$ the corresponding eigenvalues (T denotes matrix transposition). Likewise, for the signal covariance matrix let $\boldsymbol{\Sigma}_s = \mathbf{U}_s \mathbf{D}_s \mathbf{U}_s^T = \mathbf{U}_{s,r} \mathbf{D}_{s,r} \mathbf{U}_{s,r}^T$ where $\mathbf{U}_{s,r} := [\mathbf{u}_{s,1} \dots \mathbf{u}_{s,r}]$ and $\mathbf{D}_{s,r} := \text{diag}(d_{s,1}, \dots, d_{s,r})$ contain the r dominant eigenvectors and eigenvalues of $\boldsymbol{\Sigma}_s$, respectively; while $d_{s,\rho} := E[\pi_{\rho,t}^2]$. Further, let $\mathbf{U}_{s,p-r} \in \mathbb{R}^{p \times (p-r)}$ denote the matrix formed by the subspace of dimensionality $p - r$, which is perpendicular to the signal subspace $\mathbf{U}_{s,r}$. In the following, $\boldsymbol{\mu}_s$ is assumed known and subtracted from \mathbf{s}_t ; thus, without loss of generality (wlog) \mathbf{x}_t and \mathbf{s}_t are assumed zero-mean. Matrices $\boldsymbol{\Sigma}_s$ and $\boldsymbol{\Sigma}_w$ are not available, which is the case in several applications. Moreover, the following is assumed about sample covariances.

(a1) *The signal of interest \mathbf{s}_t and observation noise \mathbf{w}_t are independent across time t and identically distributed. Thus, by the strong law of large numbers the sample covariance matrix estimate $\hat{\boldsymbol{\Sigma}}_{x,n} := n^{-1} \sum_{t=1}^n \mathbf{x}_t \mathbf{x}_t^T$ converges almost surely, as $n \rightarrow \infty$, to the ensemble covariance matrix $\boldsymbol{\Sigma}_x$.*

Consider next a unitary transformation matrix $\mathbf{F} \in \mathbb{R}^{p \times p}$ to form the transformed data

$$\check{\mathbf{x}}_t := \mathbf{F} \mathbf{x}_t = \sum_{\rho=1}^r \pi_{t,\rho} \mathbf{F} \mathbf{u}_{s,\rho} + \mathbf{F} \mathbf{w}_t. \quad (2)$$

Such a mapping could represent the wavelet, Fourier, or, the discrete cosine transform (DCT). The case of interest here is when \mathbf{F} is such that the transformed eigenvectors $\check{\mathbf{u}}_{s,\rho} := \mathbf{F} \mathbf{u}_{s,\rho}$ of $\boldsymbol{\Sigma}_{\check{s}}$, where $\check{s} := \mathbf{F} \mathbf{s}$, have many entries equal to zero, i.e., $\boldsymbol{\Sigma}_{\check{s}}$ has sparse eigenvectors. One natural question is whether eigenvectors of a covariance matrix admit a sparse representation over e.g., the DCT domain. Often in bio-informatics and imaging applications the data input to the DR module have covariance matrix with sparse eigenvectors [18]. The same attribute is also present in other classes of signals. Consider for instance signal vectors comprising uncorrelated groups of entries, with each group containing correlated entries – a case where

the covariance matrix is block diagonal. In addition to block diagonal covariance matrices, the class also includes row- and/or column-permuted versions of block diagonal covariance matrices.

An example is provided next to demonstrate that Σ_s can have sparse eigenvectors. A high-resolution image taken from [1] displaying part of the Martian terrain, was split into 112 smaller non-overlapping images of size 180×256 . Each of these images was further split into 8×8 blocks. Vectors $\{s_t\}_{t=1}^{112}$ correspond to a block (here the 10th after lexicographically scanning each of the 112 sub-images) comprising entries with the same row and column indices in all 112 different sub-images. An estimate of the underlying covariance matrix of the vectorized blocks is formed using sample averaging, to obtain $\hat{\Sigma}_s := (112)^{-1} \sum_{t=1}^{112} s_t s_t^T$. The $\text{rank}(\hat{\Sigma}_s) = 15$ indicates that the 64×1 vectorized image blocks lie approximately in a linear subspace of $\mathbb{R}^{64 \times 1}$ having dimension $r = 15$. This subspace is spanned by the 15 principal eigenvectors of $\hat{\Sigma}_s$ and forms the signal subspace. To explore whether the principal eigenvectors of Σ_s have a hidden sparsity structure that could be exploited during DR Fig. 1 (left) depicts the value for each of the entries of the second principal eigenvector of $\hat{\Sigma}_s$, namely $\hat{u}_{s,2}$, versus the index number of each entry. Note that the entries of $\hat{u}_{s,2}$ exhibit a sinusoidal behavior; thus, if DCT is applied to $\hat{u}_{s,2}$ the resulting vector $\tilde{u}_{s,2} = \mathbf{F}\hat{u}_{s,2}$ has only a few DCT coefficients with large magnitude. Indeed, Fig. 1 (right) corroborates that $\hat{u}_{s,2}$ is sparse over the DCT domain. In fact, all 15 principal eigenvectors of $\hat{\Sigma}_s$ admit a sparse representation over the DCT domain as the one displayed in Fig. 1 (right). Thus, the sample covariance matrix of the transformed vectorized blocks $\tilde{s}_t = \mathbf{F}s_t$ has 15 principal eigenvectors that exhibit high degree of sparsity. Such a sparse structure is to be expected since images generated from [1] exhibit localized features (hilly terrain), which further result in sparse signal basis vectors in the DCT domain [18]. For simplicity, the original notation \mathbf{x}_t , s_t and \mathbf{w}_t will henceforth refer to the DCT transformed training data, signal of interest, and observation noise, respectively.

Aiming to compress data vector \mathbf{x} , linear DR is performed at the encoder by left-multiplying \mathbf{x} with a fat matrix $\mathbf{C} \in \mathbb{R}^{q \times p}$, where $q \leq r \leq p$. The reduced dimension q may be chosen smaller than the signal subspace dimension r , when resources are limited. Vector $\mathbf{C}\mathbf{x}$ is received at the decoder where it is left-multiplied by a tall $p \times q$ matrix \mathbf{B} to reconstruct \mathbf{s} as $\hat{\mathbf{s}} := \mathbf{B}\mathbf{C}\mathbf{x}$. Matrices \mathbf{B} and \mathbf{C} should be selected such that $\hat{\mathbf{s}} = \mathbf{B}\mathbf{C}\mathbf{x}$ forms a ‘good’ estimate of \mathbf{s} . One pair of matrices $\mathbf{B}_o, \mathbf{C}_o$, minimizing the reconstruction MSE

$$(\mathbf{B}_o, \mathbf{C}_o) \in \arg \min_{\mathbf{B}, \mathbf{C}} E[\|\mathbf{s} - \mathbf{B}\mathbf{C}\mathbf{x}\|^2] \quad (3)$$

are given as $\mathbf{C}_o = \mathbf{U}_{sx,q}^T \Sigma_{sx} \Sigma_x^{-1}$, $\mathbf{B}_o = \mathbf{U}_{sx,q}$, where $\mathbf{U}_{sx,q}$ are the q principal eigenvectors of $\Sigma_{sx} \Sigma_x^{-1} \Sigma_{sx}^T$, while the \in notation emphasizes that (3) does not have a unique optimal solution (e.g., see [5, Ch. 10]). In the absence of observation noise, ($\mathbf{x}_t = s_t$), (3) corresponds to the standard PCA,

where a possible choice for $\mathbf{B}_o, \mathbf{C}_o$ is $\mathbf{B}_o = \mathbf{U}_{s,q}$ and $\mathbf{C}_o = \mathbf{U}_{s,q}^T$ [5, Ch. 9]. Since the ensemble covariance matrices are not available; \mathbf{C}_o and \mathbf{B}_o cannot be found. The practical approach is to replace the cost in (3) with its sample-averaged version $n^{-1} \|\mathbf{S} - \mathbf{BCX}\|_F^2$, where $\mathbf{S} := [\mathbf{s}_1 \dots \mathbf{s}_n]$ and $\mathbf{X} := [\mathbf{x}_1 \dots \mathbf{x}_n]$. This would require training samples for both \mathbf{x} , and the signal of interest \mathbf{s} [5, Ch. 10]. This is impossible in the noisy setting considered here. The reduced dimension $q \leq p$ can be selected depending on the desired reduction viz reconstruction error afforded.

In the noiseless case, the optimal DR and reconstruction matrices are formed using the signal eigenvectors, that is $\mathbf{C}_o = \mathbf{B}_o^T = \mathbf{U}_{s,q}^T$. But even in the noisy case, the signal subspace $\mathbf{U}_{s,q}$ is useful for joint DR and denoising [25], [26], [28], [34]. Indeed, if $\mathbf{C} = \mathbf{B}^T = \mathbf{U}_{s,q}^T$, then it is easy to see that $J_{\text{rec}}(\mathbf{B}, \mathbf{C}) := E[\|\mathbf{s} - \mathbf{BCx}\|_2^2] = \text{tr}(\Sigma_w) - \text{tr}(\mathbf{U}_{s,p-q}^T \Sigma_w \mathbf{U}_{s,p-q}) < \text{tr}(\Sigma_w)$ when $q = r < p$. Thus, projection of \mathbf{x}_t onto the signal subspace not only achieves DR but also reduces noise effects. The question of course is how to form an estimate for $\mathbf{U}_{s,q}$ when Σ_x and Σ_w are unknown. Existing signal subspace estimators assume either that i) the noise is white, namely $\Sigma_w = \sigma_w^2 \mathbf{I}_p$ for which $\mathbf{U}_x = \mathbf{U}_s$ (\mathbf{I}_p denotes the $p \times p$ identity matrix); or, ii) the Σ_w is known or can be estimated via sample-averaging [25], [26], [28], [34]. In the setting here these assumptions are not needed.

Based on training data $\{\mathbf{x}_t\}_{t=1}^n$, the major goal is to exploit the sparsity present in the eigenvectors of Σ_s in order to derive estimates $\hat{\mathbf{B}}, \hat{\mathbf{C}}$ for the signal subspace $\mathbf{U}_{s,q}$, thereby achieving a better trade-off between the reduced-dimension q and the MSE cost $J_{\text{rec}}(\mathbf{B}, \mathbf{C})$ than existing alternatives [9], [20], [33], [36]. Towards this end, a novel sparsity-aware signal subspace estimator is developed in the next section. Since the majority of data processing and communication systems are digital, this sparsity-exploiting DR step will be followed by a sparsity-cognizant VQ step in order to develop a sparsity-aware transform coding (SATC) scheme for recovering \mathbf{s}_t based on quantized DR data.

III. SPARSE PRINCIPAL COMPONENT ANALYSIS

Recall that the standard PCA determines DR and reconstruction matrices $\hat{\mathbf{C}}_o$ and $\hat{\mathbf{B}}_o$ by minimizing the sample-based cost $\hat{J}_{\text{rec}}(\mathbf{B}, \mathbf{C}) := n^{-1} \|\mathbf{X} - \mathbf{BCX}\|_F^2$. One possible minimizer for the latter is $\hat{\mathbf{B}}_o = \hat{\mathbf{C}}_o^T = \hat{\mathbf{U}}_{x,q}$, where $\hat{\mathbf{U}}_{x,q}$ comprises the q dominant eigenvectors of $\hat{\Sigma}_x$. In the noiseless case it holds that $\hat{\Sigma}_x = \hat{\Sigma}_s$, from which it follows that $\hat{\mathbf{U}}_{x,q} = \hat{\mathbf{U}}_{s,q}$. However, the q -dominant eigenvectors of Σ_x do not coincide with $\hat{\mathbf{U}}_{s,q}$ when the additive noise is colored ($\Sigma_w \neq \sigma_w^2 \mathbf{I}_p$). In this case, standard PCA is not expected to estimate reliably the signal subspace.

A. An ℓ_1 -regularized formulation

Here the standard PCA formulation is enhanced by exploiting the sparsity present in $\mathbf{U}_{s,r}$. Prompted by Lasso-based sparse regression and PCA approaches [9], [20], [29], [33], [36], the quadratic cost of standard PCA is regularized with the ℓ_1 -norm of the unknowns to effect sparsity. Specifically, $\mathbf{B}_o = \mathbf{C}_o^T = \mathbf{U}_{s,q}$ could be estimated as

$$\arg \min_{\mathbf{B}, \mathbf{C}} n^{-1} \|\mathbf{X} - \mathbf{B}\mathbf{C}\mathbf{X}\|_F^2 + \sum_{\rho=1}^q \sum_{j=1}^p \lambda_{\rho} (|\mathbf{C}(\rho, j)| + |\mathbf{B}(j, \rho)|), \text{ s. to } \mathbf{B} = \mathbf{C}^T \quad (4)$$

which promotes sparsity in $\mathbf{U}_{s,r}$. However, the constraint $\mathbf{B} = \mathbf{C}^T$ leads to a nonconvex problem that cannot be solved efficiently. This motivates the following ‘relaxed’ version of (4), where the wanted matrices are obtained as

$$(\hat{\mathbf{B}}, \hat{\mathbf{C}}) \in \arg \min_{\mathbf{B}, \mathbf{C}} n^{-1} \|\mathbf{X} - \mathbf{B}\mathbf{C}\mathbf{X}\|_F^2 + \sum_{\rho=1}^q \sum_{j=1}^p \lambda_{\rho} (|\mathbf{C}(\rho, j)| + |\mathbf{B}(j, \rho)|) + \mu \|\mathbf{B} - \mathbf{C}^T\|_F^2 \quad (5)$$

using efficient coordinate descent solvers. Note that $q \leq r$, since the dimensionality of the signal subspace may not be known. Moreover, $\{\lambda_{\rho}\}_{\rho=1}^q$ are nonnegative constants controlling the sparsity of \mathbf{B} and \mathbf{C} . Indeed, the larger λ_{ρ} ’s are chosen, the closer the entries of \mathbf{B} and \mathbf{C} are driven to the origin. Taking into account that the ‘clairvoyant’ compression and reconstruction matrices satisfy $\mathbf{B}_o = \mathbf{C}_o^T = \mathbf{U}_{s,q}$, the term $\mu \|\mathbf{B} - \mathbf{C}^T\|_F^2$ ensures that $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}^T$ stay close. Although \mathbf{B}_o and \mathbf{C}_o are orthonormal, \mathbf{B} and \mathbf{C} are not constrained to be orthonormal in (5) because orthonormality constraints of the form $\mathbf{B}^T \mathbf{B} = \mathbf{I}$ and $\mathbf{C}\mathbf{C}^T = \mathbf{I}$ are nonconvex. With such constraints present, efficient coordinate descent algorithms converging to a stationary point cannot be guaranteed.

Remark 1: The minimization problem in (5) resembles the sparse PCA formulation proposed in [36]. However, the approach followed in [36] imposes sparsity *only* on \mathbf{C} , while it forces matrix \mathbf{B} to be orthonormal. The latter constraint mitigates scaling issues (otherwise \mathbf{C} could be made arbitrarily small by counter-scaling \mathbf{B}), but is otherwise not necessarily well-motivated. Without effecting sparsity in \mathbf{B} , the formulation in [36] does not fully exploit the sparsity present in the eigenvectors of Σ_s , which further results in sparse clairvoyant matrices \mathbf{C}_o and \mathbf{B}_o in the absence of noise. Notwithstanding, (5) combines the reconstruction error $n^{-1} \|\mathbf{X} - \mathbf{B}\mathbf{C}\mathbf{X}\|_F^2$ with regularization terms that impose sparsity to *both* \mathbf{B} and \mathbf{C} . Even though the ℓ_1 -norm of \mathbf{C} together with $\|\mathbf{B} - \mathbf{C}^T\|_F^2$ suffice to prevent scaling issues, the ℓ_1 -norm of \mathbf{B} is still needed to ensure sparsity in the entries of $\hat{\mathbf{B}}$.

B. Block Coordinate Descent Algorithm

The minimization problem in (5) is nonconvex with respect to (wrt) both \mathbf{B} and \mathbf{C} . This challenge will be bypassed by constructing an iterative solver. Relying on block coordinate descent (see e.g., [3, pg.

160]) the cost in (5) will be iteratively minimized wrt \mathbf{B} (or \mathbf{C}), while keeping matrix \mathbf{C} (or \mathbf{B}) fixed.

Specifically, given the matrix $\hat{\mathbf{B}}_{\tau-1}$ at the end of iteration $\tau-1$, an updated estimate of \mathbf{C}_o at iteration τ can be formed by solving the minimization problem [cost in (5) has been scaled with n]

$$\hat{\mathbf{C}}_\tau = \arg \min_{\mathbf{C}} \|\mathbf{X} - \hat{\mathbf{B}}_{\tau-1} \mathbf{C} \mathbf{X}\|_F^2 + \sum_{\rho=1}^q \lambda_\rho \|\mathbf{C}_{\rho:}^T\|_1 + \mu \|\mathbf{C} - \hat{\mathbf{B}}_{\tau-1}^T\|_F^2 \quad (6)$$

where $\mathbf{C}_{\rho:}^T$ denotes the ρ th row of \mathbf{C} , while n is absorbed in μ and λ_ρ . After straightforward manipulations (6) can be equivalently reformulated as

$$\hat{\mathbf{C}}_\tau = \arg \min_{\mathbf{C}} \text{tr}(\mathbf{X}^T \mathbf{C}^T \hat{\mathbf{B}}_{\tau-1}^T \hat{\mathbf{B}}_{\tau-1} \mathbf{C} \mathbf{X} - 2 \mathbf{X}^T \mathbf{C}^T \hat{\mathbf{B}}_{\tau-1}^T \mathbf{X}) + \sum_{\rho=1}^q [\lambda_\rho \|\mathbf{C}_{\rho:}^T\|_1 + \mu \|\mathbf{C}_{\rho:}^T\|_2^2 - 2\mu \mathbf{C}_{\rho:}^T \hat{\mathbf{B}}_{\tau-1,:\rho}] \quad (7)$$

where $\hat{\mathbf{B}}_{\tau-1,:\rho}$ corresponds to the ρ th column of $\hat{\mathbf{B}}_{\tau-1}$. After introducing some auxiliary variables $\{t_{\rho,j}\}_{j=1,\rho=1}^{p,q}$, the optimization problem in (7) can be equivalently rewritten as a convex optimization problem that has i) a cost given by $\text{tr}(\mathbf{X}^T \mathbf{C}^T \hat{\mathbf{B}}_{\tau-1}^T \hat{\mathbf{B}}_{\tau-1} \mathbf{C} \mathbf{X}) - 2\text{tr}(\mathbf{X}^T \mathbf{C}^T \hat{\mathbf{B}}_{\tau-1}^T \mathbf{X}) + \sum_{\rho=1}^q \lambda_\rho \sum_{j=1}^p t_{\rho,j} + \mu \sum_{\rho=1}^q \|\mathbf{C}_{\rho:}^T\|_2^2 - 2\mu \sum_{\rho=1}^q \mathbf{C}_{\rho:}^T \hat{\mathbf{B}}_{\tau-1,:\rho}$; and ii) a constraint set formed by the inequalities $\{|\mathbf{C}(\rho, j)| \leq t_{\rho,j}\}_{j=1,\rho=1}^{p,q}$. This constrained minimization problem can be solved using an interior point method [4].

Given the most recent DR update $\hat{\mathbf{C}}_\tau$, an updated estimate of the reconstruction matrix \mathbf{B}_o is obtained as

$$\hat{\mathbf{B}}_\tau = \arg \min_{\mathbf{B}} \|\mathbf{X} - \mathbf{B} \hat{\mathbf{C}}_\tau \mathbf{X}\|_F^2 + \sum_{\rho=1}^q \lambda_\rho \|\mathbf{B}_{\cdot,\rho}\|_1 + \mu \|\mathbf{B} - \hat{\mathbf{C}}_\tau^T\|_F^2. \quad (8)$$

The minimization problem in (8) can be split into the following p subproblems:

$$\begin{aligned} \hat{\mathbf{B}}_{\tau,j} &= \arg \min_{\mathbf{B}_{j:}} \|\mathbf{X}_{j:} - \mathbf{B}_{j:}^T \hat{\mathbf{C}}_\tau \mathbf{X}\|_2^2 + \mu \|\mathbf{B}_{j:}^T - (\hat{\mathbf{C}}_{\tau,j})^T\|_2^2 + \sum_{\rho=1}^q \lambda_\rho |\mathbf{B}(j, \rho)| \\ &= \arg \min_{\mathbf{B}_{j:}} \|[\mathbf{X}_{j:}, \mu^{1/2}(\hat{\mathbf{C}}_{\tau,j})^T] - \mathbf{B}_{j:}^T [\hat{\mathbf{C}}_\tau \mathbf{X}, \mu^{1/2} \mathbf{I}_q]\|_2^2 + \sum_{\rho=1}^q \lambda_\rho |\mathbf{B}(j, \rho)|, \quad j = 1, \dots, p \end{aligned} \quad (9)$$

where $\mathbf{X}_{j:}$ denotes the j th row of \mathbf{X} .

Notice that (9) corresponds to a Lasso problem that can be solved efficiently using e.g., the LARS algorithm [10]. The proposed block coordinate descent (BCD-) S-PCA algorithm yields iterates $\hat{\mathbf{B}}_\tau$ and $\hat{\mathbf{C}}_\tau$ that converge, as $\tau \rightarrow \infty$, at least to a stationary point of the cost in (5)- a fact established using the results in e.g., [30, Thm. 4.1] (see also arguments in Apdx. B). The BCD-SPCA scheme is tabulated as Algorithm 1.

C. Efficient SPCA Solver

Relying on the BCD-SPCA algorithm of the previous section, an element-wise coordinate descent algorithm is developed here to numerically solve (5) with reduced computational complexity. Specifically,

Algorithm 1 : BCD-SPCA

Initialize $\hat{\mathbf{B}}_0 = \hat{\mathbf{C}}_0^T = \hat{\mathbf{U}}_{s,q}$, where $\hat{\mathbf{U}}_{s,q}$ is the signal subspace estimate found by standard PCA.

for $\tau = 1, \dots$ **do**

Find $\hat{\mathbf{C}}_\tau$ by solving (7).

Find $\hat{\mathbf{B}}_\tau$ by solving (9) for $j = 1, \dots, p$.

If $|\text{Cost}_\tau - \text{Cost}_{\tau-1}| < \epsilon$ for a prescribed tolerance ϵ **then break**

end for

the cost in (5) is iteratively minimized wrt an entry of either \mathbf{B} or \mathbf{C} , while keeping the remaining entries fixed. One coordinate descent iteration involves updating all the entries of matrices \mathbf{B} and \mathbf{C} .

Given iterates $\hat{\mathbf{B}}_{\tau-1}$ and $\hat{\mathbf{C}}_{\tau-1}$, the next steps describe how the entries of $\hat{\mathbf{C}}_\tau$ and $\hat{\mathbf{B}}_\tau$ are formed. Let \otimes denote the Kronecker product, and $\text{vec}(\mathbf{C})$ the $qp \times 1$ vector obtained after stacking the p columns of \mathbf{C} . Using the property $\text{vec}(\hat{\mathbf{B}}_{\tau-1} \mathbf{C} \mathbf{X}) = (\mathbf{X}^T \otimes \hat{\mathbf{B}}_{\tau-1}) \text{vec}(\mathbf{C})$, the cost in (5) after setting $\mathbf{B} = \hat{\mathbf{B}}_{\tau-1}$ can be re-expressed as

$$\|\text{vec}(\mathbf{X}) - (\mathbf{X}^T \otimes \hat{\mathbf{B}}_{\tau-1}) \text{vec}(\mathbf{C})\|_2^2 + \sum_{\rho=1}^q \lambda_\rho (\|\mathbf{C}_\rho^T\|_1 + \|\hat{\mathbf{B}}_{\tau-1,:\rho}\|_1) + \mu \|\hat{\mathbf{B}}_{\tau-1} - \mathbf{C}^T\|_F^2. \quad (10)$$

Next, we show how to form $\hat{\mathbf{C}}_\tau(\rho, j)$ at iteration τ , based on the most up-to-date values of \mathbf{B} and $\mathbf{c}_v := \text{vec}(\mathbf{C})$, namely $\hat{\mathbf{B}}_{\tau-1}$, $\{\hat{\mathbf{c}}_{v,\tau-1}(m)\}_{m=(j-1)q+\rho+1}^{qp}$ and $\{\hat{\mathbf{c}}_{v,\tau}(m)\}_{m=1}^{(j-1)q+\rho-1}$. It follows from (10) that $\hat{\mathbf{C}}_\tau(\rho, j) \equiv \hat{\mathbf{c}}_{v,\tau}((j-1)q + \rho)$, for $\rho = 1, \dots, q$ and $j = 1, \dots, p$, can be determined as

$$\hat{\mathbf{C}}_\tau(\rho, j) = \arg \min_c \|\chi_{\tau,\rho,j} - c \hat{\mathbf{X}}_{B_{\tau-1},:(j-1)q+\rho}\|_2^2 + \mu(c - \hat{\mathbf{B}}_{\tau-1}(j, \rho))^2 + \lambda_\rho |c| \quad (11)$$

where $\hat{\mathbf{X}}_{B_{\tau-1}} := \mathbf{X}^T \otimes \hat{\mathbf{B}}_{\tau-1}$, $\chi_{\tau,\rho,j} := \text{vec}(\mathbf{X}) - \sum_{m=1}^{(j-1)q+\rho-1} \hat{\mathbf{c}}_{v,\tau}(m) \hat{\mathbf{X}}_{B_{\tau-1},:m} - \sum_{m=(j-1)q+\rho+1}^{qp} \hat{\mathbf{c}}_{v,\tau-1}(m) \hat{\mathbf{X}}_{B_{\tau-1},:m}$, and $\hat{\mathbf{X}}_{B_{\tau-1},:m}$ corresponds to the m th column of $\hat{\mathbf{X}}_{B_{\tau-1}}$. Interestingly, the minimization in (11) corresponds to a sparse regression (Lasso) problem involving a scalar. The latter admits a closed-form solution which is given in the next Lemma (see Apdx. A for the proof).

Lemma 1 *The optimal solution of the minimization problem*

$$\hat{c} = \arg \min_c \|\chi - c\mathbf{h}\|_2^2 + \mu(c - \hat{b})^2 + \lambda|c| \quad (12)$$

where \mathbf{y} and \mathbf{h} are column vectors and \hat{b} is a scalar constant, is given by

$$\hat{c} = \text{sgn}\left(\chi^T \mathbf{h} + \mu \hat{b}\right) \times \left(\left| \frac{\chi^T \mathbf{h} + \mu \hat{b}}{\|\mathbf{h}\|_2^2 + \mu} \right| - \frac{\lambda}{2\|\mathbf{h}\|_2^2 + 2\mu} \right)_+.$$

Based on Lemma 1 one can readily deduce that the optimal solution of (11) is given by

$$\hat{\mathbf{C}}_\tau(\rho, j) = \text{sgn} \left((\mathbf{x}_{\tau, \rho, j})^T \hat{\mathbf{X}}_{B_{\tau-1}, : (j-1)q + \rho} + \mu \hat{\mathbf{B}}_{\tau-1}(j, \rho) \right) \times \left(\left| \frac{(\mathbf{x}_{\tau, \rho, j})^T \hat{\mathbf{X}}_{B_{\tau-1}, : (j-1)q + \rho} + \mu \hat{\mathbf{B}}_{\tau-1}(j, \rho)}{\|\hat{\mathbf{X}}_{B_{\tau-1}, : (j-1)q + \rho}\|_2^2 + \mu} \right| - \frac{\lambda_\rho}{2\|\hat{\mathbf{X}}_{B_{\tau-1}, : (j-1)q + \rho}\|_2^2 + 2\mu} \right)_+ \quad (13)$$

where $(\cdot)_+ := \max(\cdot, 0)$.

Similarly, starting from the minimization problem in (9) and applying an element-wise coordinate descent approach an update for the $\mathbf{B}(j, \rho)$ can be obtained as

$$\hat{\mathbf{B}}_\tau(j, \rho) = \arg \min_{\mathbf{b}} \|\psi_{\tau, j, \rho} - \mathbf{b} \hat{\mathbf{X}}_{C_\tau, : \rho}\|_2^2 + \mu(b - \hat{\mathbf{C}}_\tau(\rho, j))^2 + \lambda_\rho |b|, \quad \rho = 1, \dots, q \quad (14)$$

where $\hat{\mathbf{X}}_{C_\tau} := \mathbf{X}^T \hat{\mathbf{C}}_\tau^T$, $\psi_{\tau, j, \rho} := \mathbf{X}_{:,j} - \sum_{l=1}^{\rho-1} \hat{\mathbf{B}}_\tau(j, l) \hat{\mathbf{X}}_{C_\tau, : l} - \sum_{l=\rho+1}^q \hat{\mathbf{B}}_{\tau-1}(j, l) \hat{\mathbf{X}}_{C_\tau, : l}$, and $\hat{\mathbf{X}}_{C_\tau, : l}$ denotes the l th column of $\hat{\mathbf{X}}_{C_\tau}$, while $\mathbf{X}_{:,j}$ refers to the j th column of \mathbf{X} . The optimal solution of the minimization problem in (14) is given by

$$\hat{\mathbf{B}}_\tau(j, \rho) = \text{sgn} \left((\psi_{\tau, j, \rho})^T \hat{\mathbf{X}}_{C_\tau, : \rho} + \mu \hat{\mathbf{C}}_\tau(\rho, j) \right) \times \left(\left| \frac{(\psi_{\tau, j, \rho})^T \hat{\mathbf{X}}_{C_\tau, : \rho} + \mu \hat{\mathbf{C}}_\tau(\rho, j)}{\|\hat{\mathbf{X}}_{C_\tau, : \rho}\|_2^2 + \mu} \right| - \frac{\lambda_\rho}{2\|\hat{\mathbf{X}}_{C_\tau, : \rho}\|_2^2 + 2\mu} \right)_+ \quad (15)$$

Note that iteration τ involves minimizing (5) wrt to each entry of \mathbf{C} or \mathbf{B} while fixing the rest. It is shown in Appendix B that the computationally efficient coordinate descent (ECD)-SPCA scheme converges, as $\tau \rightarrow \infty$, at least to a stationary point of the cost in (5) when the entries of \mathbf{X} are finite. The proof relies on [30, Thm. 4.1]. Using arguments similar to those in Appendix B, it can be shown that BCD-SPCA converges too. A stationary point for the nondifferentiable cost here is defined as one whose lower directional derivative is nonnegative toward any possible direction [30, Sections 1 and 3], meaning that the cost cannot *decrease* when moving along any possible direction around and close to a stationary point. A strictly positive μ ensures that the minimization problems in (11) and (14) are *strictly* convex with respect to either \mathbf{C} or \mathbf{B} . This guarantees that the minimization problems in (11) or (14) have a unique minimum per iteration, which in turn implies that the ECD-SPCA algorithm converges to a stationary point. If $\mu = 0$, the proposed algorithms may not converge. This can also be seen from the updating recursions (13) and (15). If $\mu = 0$ and at a certain iteration one of the matrices is zero, say $\hat{\mathbf{C}}_\tau$, this may cause the entries of $\hat{\mathbf{B}}_\tau$ to diverge. Similar comments apply in BCD-SPCA.

The optimal solution in (13) and (15) can be determined in closed form at computational complexity $\mathcal{O}(np)$. With $2qp$ entries in \mathbf{C} and \mathbf{B} , the total complexity for completing a full coordinate descent iteration is $\mathcal{O}(nqp^2)$. The ECD-SPCA scheme is tabulated as Algorithm 2.

Note that the sparsity coefficient λ_ρ is common to both terms $\|\mathbf{C}_{\rho,:}^T\|_1$ and $\|\mathbf{B}_{:, \rho}\|_1$. This together with the explicit dissimilarity penalty in (5) force the estimates obtained via ECD-SPCA (denoted $\hat{\mathbf{B}}_\tau$ and $\hat{\mathbf{C}}_\tau^T$)

Algorithm 2 : ECD-SPCA

Initialize $\hat{\mathbf{B}}_0 = \hat{\mathbf{C}}_0^T = \hat{\mathbf{U}}_{s,q}$, where $\hat{\mathbf{U}}_{s,q}$ is the signal subspace estimate found by standard PCA.

for $\tau = 1, \dots$ **do**

For $j = 1, \dots, p$ and $\rho = 1, \dots, q$ determine $\hat{\mathbf{C}}_\tau(\rho, j)$ using (13).

For $j = 1, \dots, p$ and $\rho = 1, \dots, q$ determine $\hat{\mathbf{B}}_\tau(j, \rho)$ using (15).

If $|\text{Cost}_\tau - \text{Cost}_{\tau-1}| < \epsilon$ for a prescribed tolerance ϵ **then break**

end for

to be approximately equal for sufficiently large τ . The latter equality requirement is further enforced in the BCD- (or ECD-) SPCA scheme by selecting μ sufficiently large, e.g., in our setting $\mu = 100$. The ideal \mathbf{B}_o and \mathbf{C}_o^T are orthonormal and equal in the noiseless case; thus, the same properties are imposed to $\hat{\mathbf{B}}_\tau$ and $\hat{\mathbf{C}}_\tau^T$. Towards this end, we: i) pick one of the matrices $\hat{\mathbf{B}}_\tau$ and $\hat{\mathbf{C}}_\tau^T$, say the latter; ii) extract, using SVD, an orthonormal basis $\hat{\mathbf{U}}_s \in \mathbb{R}^{p \times q}$ spanning the range space of $\hat{\mathbf{C}}_\tau^T$; and iii) form the compression and reconstruction matrices $\hat{\mathbf{B}}_\tau = \hat{\mathbf{C}}_\tau^T = \hat{\mathbf{U}}_s$. Thus, the dimensionality of each acquired vector \mathbf{x} is reduced at the encoder using the linear operator $\hat{\mathbf{U}}_s$, while at the decoder the signal of interest is reconstructed as $\hat{\mathbf{s}} = \hat{\mathbf{U}}_s(\hat{\mathbf{U}}_s^T \mathbf{x}) = (\hat{\mathbf{C}}_\tau^T \hat{\mathbf{C}}_\tau)^\dagger \hat{\mathbf{C}}_\tau^T (\hat{\mathbf{C}}_\tau \mathbf{x})$, where the symbol \dagger denotes matrix pseudoinverse.

D. Tuning the sparsity-controlling coefficients

Up to now the sparsity-controlling coefficients were assumed given. A cross-validation (CV) based algorithm is derived in this section to select $\{\lambda_\rho\}_{\rho=1}^q$ with the objective to find estimates $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ leading to good reconstruction performance, i.e., small $J_{\text{rec}}(\hat{\mathbf{B}}, \hat{\mathbf{C}})$. The CV scheme is developed for the noiseless case ($\mathbf{x}_t = \mathbf{s}_t$).

Consider the M -fold CV scheme in [17, pg. 241-249], for selecting $\{\lambda_\rho\}_{\rho=1}^q$ from a q -dimensional grid $\mathcal{L} := \{\lambda_1^1, \dots, \lambda_1^J\} \times \dots \times \{\lambda_q^1, \dots, \lambda_q^J\}$, where $\{0 < \lambda_\rho^1 < \dots < \lambda_\rho^J\}_{\rho=1}^q$ denote the candidate values, and \times denotes Cartesian product. The training data set $\{\mathbf{x}_t\}_{t=1}^n$ is divided into M nonoverlapping subsets $\{\mathbf{X}_m\}_{m=1}^M$. Let $\hat{\mathbf{B}}_{-m}(\{\lambda_\rho\}_{\rho=1}^q)$ and $\hat{\mathbf{C}}_{-m}(\{\lambda_\rho\}_{\rho=1}^q)$ denote the estimates obtained via BCD-SPCA (or ECD-SPCA) when using all the training data but those in \mathbf{X}_m , for fixed values of the sparsity-controlling coefficients. The next step is to evaluate an estimate of the reconstruction error using \mathbf{X}_m , i.e., form the reconstruction error estimate $\hat{J}_{\text{rec}}^m(\{\lambda_\rho\}_{\rho=1}^q) := |\mathbf{X}_m|^{-1} \|\mathbf{X}_m - \hat{\mathbf{B}}_{-m} \hat{\mathbf{C}}_{-m} \mathbf{X}_m\|_F^2$, where $|\mathbf{X}_m|$ indicates the cardinality of \mathbf{X}_m . A sample-based estimate of the reconstruction MSE can be found as

$$\hat{J}_{\text{rec}}(\{\lambda_\rho\}_{\rho=1}^q) = n^{-1} \sum_{t=1}^n \hat{J}_{\text{rec}}^{m(t)}(\{\lambda_\rho\}_{\rho=1}^q) \quad (16)$$

where $m(t)$ denotes the partition index in which \mathbf{x}_t is included during the CV process.

Using (16), the desired sparsity-controlling coefficients are selected as

$$(\{\hat{\lambda}_\rho\}_{\rho=1}^q) := \arg \min_{\{\lambda_\rho\} \in \mathcal{L}} \hat{J}_{\text{rec}}(\{\lambda_\rho\}_{\rho=1}^q). \quad (17)$$

The minimization (17) is carried out using exhaustive search over the grid \mathcal{L} . Fig. 2 (right) shows how the reconstruction error $\hat{J}_{\text{rec}}(\cdot)$ is affected by the sparsity controlling coefficients. A simplified scenario is considered here with $\{\lambda_\rho = \lambda\}_{\rho=1}^q$, with $p = 14$, and $q = r = 2$. Matrix $\Sigma_x = \Sigma_s$ is constructed so that 80% of the \mathbf{U}_s entries are zero. The black bars in Fig. 2 (right) quantify the standard error associated with each reconstruction MSE estimate in the red CV curve and its amplitude is calculated as

$$\text{Std}(\lambda) = \sqrt{n^{-1}} \sqrt{(n-1)^{-1} \sum_{t=1}^n [\hat{J}_{\text{rec}}^{m(t)}(\lambda) - \hat{J}_{\text{rec}}(\lambda)]^2}. \quad (18)$$

When $\lambda < 0.1$ the reconstruction MSE remains almost constant and equal to the one achieved by standard PCA ($\lambda = 0$). If $\lambda > 10^{0.5}$, the reconstruction MSE increases and reaches a maximum equal to the trace of Σ_x (the DR and reconstruction matrices equal zero). The minimum MSE is achieved for $\lambda \approx 1$. Note that λ_ρ 's have so far been selected for a fixed value of μ . Recall that μ controls the dissimilarity, of $\hat{\mathbf{C}}^T$ with $\hat{\mathbf{B}}$, thus a relatively large value ($\mu = 100$ was used in the simulations) suffices to ensure that $\hat{\mathbf{C}}$ and $\hat{\mathbf{B}}$ stay close. Of course, the higher μ is the closer $\hat{\mathbf{C}}^T$ and $\hat{\mathbf{B}}$ will be.

IV. S-PCA PROPERTIES

In this section sufficiently many training data are assumed available ($n \rightarrow \infty$), to allow analysis based on the ensemble covariance $\Sigma_x = \Sigma_s + \Sigma_w$. Recall that (a1) ensures a.s. convergence of the sample-based cost in (5) to its ensemble counterpart

$$(\mathbf{B}_e, \mathbf{C}_e) \in \arg \min_{\mathbf{B}, \mathbf{C}} \text{tr}[(\mathbf{I} - \mathbf{B}\mathbf{C})\Sigma_x(\mathbf{I} - \mathbf{B}\mathbf{C})^T] + \sum_{\rho=1}^q \lambda_\rho (\|\mathbf{C}_{\rho:}^T\|_1 + \|\mathbf{B}_{:\rho}\|_1) + \mu \|\mathbf{B} - \mathbf{C}^T\|_F^2. \quad (19)$$

Interestingly, it will turn out that even in the presence of colored noise \mathbf{w}_t , the solution pair $\mathbf{B}_e, \mathbf{C}_e$ can recover the support of the columns of the signal subspace $\mathbf{U}_{s,r}$, or at least part of it, as long as the noise power in \mathbf{x}_t is sufficiently small.

A. Support recovery in colored noise

In this section entries of \mathbf{C}_e (or \mathbf{B}_e) will be considered nonzero only if their corresponding magnitudes exceed an arbitrarily small threshold $\delta > 0$. Under this condition it will be demonstrated next that for properly selected μ and λ_ρ , S-PCA assigns the (non)zero entries in \mathbf{C}_e consistently with the support of the

columns of $\mathbf{U}_{s,r}$. This means that the S-PCA formulation is meaningful because it does not assign entries of \mathbf{C}_e (or \mathbf{B}_e) arbitrarily, but takes into account where the (non)zero entries of $\mathbf{U}_{s,r}$ are. Interestingly, this will hold even when colored noise is present, as long as its variance is properly upper bounded.

To proceed, let Σ_s be a permuted version of a block diagonal matrix. Specifically:

(a2) *The entries of \mathbf{s} can be partitioned into groups $\mathcal{G}_1, \dots, \mathcal{G}_K$, so that entries with indices in the same group are allowed to be correlated but entries with indices in different groups are uncorrelated; i.e., if $j \in \mathcal{G}_k$ and $j' \in \mathcal{G}_{k'}$, then $E[\mathbf{s}(j)\mathbf{s}(j')] = 0$ for $k \neq k'$. Moreover, let $\{\mathcal{G}_k\}_{k=1}^K$ have the same cardinality that is equal to G . Using the proper permutation matrix \mathbf{P} , these groups can be made contiguous; hence, the vector $\mathbf{s}_P := \mathbf{P}\mathbf{s} = [\mathbf{s}_{\mathcal{G}_1} \dots \mathbf{s}_{\mathcal{G}_K}]^T$ has covariance matrix with block diagonal structure, that is $\mathbf{P}\Sigma_s\mathbf{P}^T = \text{bdiag}(\Sigma_{s_{\mathcal{G}_1}}, \dots, \Sigma_{s_{\mathcal{G}_K}})$. This implies that the eigenvector matrix $\tilde{\mathbf{U}}_s$ of $\mathbf{P}\Sigma_s\mathbf{P}^T$ is block diagonal and sparse. Since $\mathbf{U}_s = \mathbf{P}^T\tilde{\mathbf{U}}_s$ and \mathbf{P} is a permutation matrix, it follows that \mathbf{U}_s is also sparse.*

The block-diagonal structure under (a2) emerges when \mathbf{s}_t corresponds e.g., to a random field in which the groups $\{\mathcal{G}_k\}_{k=1}^K$ correspond to different regions affected by groups of uncorrelated sources. Each of the sub-vectors $\mathbf{s}_{\mathcal{G}_k}$ in \mathbf{s}_P contains sources affecting a certain region in the field and are uncorrelated from the other sources present in \mathbf{s}_P . It is worth stressing that (a2) does not prevent applicability of the ECD-SPCA (or BCD-SPCA) algorithm, but it is introduced here only to assess its asymptotic performance. Before stating the result proved in Appendix C, let $\mathbf{v}(\mathcal{F})$ denote the entries of vector \mathbf{v} with indices belonging to the set \mathcal{F} . Further, let $\mathcal{S}(\mathbf{v})$ denote the support of \mathbf{v} , i.e., the set of indices of the nonzero entries of \mathbf{v} .

Proposition 1 *Let $\Sigma_x = \Sigma_s + \Sigma_w$, with Σ_s satisfying (a2). Further, assume that the spectral radius of Σ_w , namely $d_{\max}(\Sigma_w)$, satisfies $d_{\max}(\Sigma_w) < \Delta(\Sigma_s)$, where $\Delta(\Sigma_s) > 0$ is a function of Σ_s . If $\{\lambda_\rho = \lambda\}_{\rho=1}^q$ are selected such that $\|\mathbf{C}_{e,\rho}\|_0 \geq 2$, then for any arbitrarily small $\delta > 0$ there exists a μ_o such that for any $\mu \geq \mu_o$ the minimization in (19) admits an optimal solution satisfying*

$$\|\mathbf{C}_{e,\rho}^T(\bar{\mathcal{S}}_{i_\rho})\|_1 \leq \delta, \text{ and } \|\mathbf{C}_{e,\rho}^T(\mathcal{S}_{i_\rho})\|_1 \geq \xi(\lambda_\rho) > 0 \quad (20)$$

$$\|\mathbf{B}_{e,:\rho}(\bar{\mathcal{S}}_{i_\rho})\|_1 \leq \delta, \text{ and } \|\mathbf{B}_{e,:\rho}(\mathcal{S}_{i_\rho})\|_1 \geq \xi(\lambda_\rho) > 0, \text{ for } \rho = 1, \dots, q \quad (21)$$

where $\bar{\mathcal{S}}_{i_\rho}$ is the complement of the support \mathcal{S}_{i_ρ} of \mathbf{u}_{s,i_ρ} , while $\{i_1, \dots, i_q\} \subseteq \{1, \dots, r\}$. The constant $\xi_\rho(\lambda_\rho)$ depends only on λ_ρ and is strictly positive for a finite λ_ρ .

Prop. 1 asserts that for n sufficiently large, S-PCA has an optimal solution $(\mathbf{B}_e, \mathbf{C}_e)$ whose support is a subset of the true support of $\{\mathbf{u}_{s,i_{\rho=1}}\}_{i=1}^q$ even in the presence of colored noise. This is possible since for the ρ th row of \mathbf{C}_e , there is a corresponding i_ρ column of matrix $\mathbf{U}_{s,r}$ such that $\|\mathbf{C}_{e,\rho}^T(\bar{\mathcal{S}}_{i_\rho})\|_1 \leq \delta$ for arbitrarily small δ , while $\|\mathbf{C}_{e,\rho}^T(\mathcal{S}_{i_\rho})\|_1 \geq \xi_\rho(\lambda_\rho) > \delta \geq 0$ (strictly positive). Thus, all the nonzero entries of $\mathbf{C}_{e,\rho}^T$ with magnitude exceeding δ will have indices in $\mathcal{S}_{i_\rho} := \text{support}(\mathbf{u}_{s,i_\rho})$. This happens since: i)

$\|\mathbf{C}_{e,\rho}^T(\bar{\mathcal{S}}_{i_\rho})\|_1$ can be made arbitrarily small, thus all entries of $\mathbf{C}_{e,\rho}^T$ with indices in $\bar{\mathcal{S}}_{i_\rho}$ can be driven arbitrarily close (δ -close) to zero by controlling μ ; and ii) $\|\mathbf{C}_{e,\rho}^T(\mathcal{S}_{i_\rho})\|_1$ is strictly positive with $\xi_\rho(\lambda_\rho) > \delta$, thus some of the entries of $\mathbf{C}_{e,\rho}^T$ with indices in \mathcal{S}_{i_ρ} must have magnitude greater than δ . The number of nonzero entries in $\mathbf{C}_{e,\rho}^T(\mathcal{S}_{i_\rho})$ is determined by λ_ρ . Thus, if λ_ρ is selected such that $\|\mathbf{C}_{e,\rho}^T\|_0 = \|\mathbf{u}_{s,i_\rho}\|_0$, then recovery of the whole support \mathcal{S}_{i_ρ} is ensured.

Remark 2: It should be clarified that the vectors $\{\mathbf{u}_{s,i_\rho}\}_{\rho=1}^q$ in Prop. 1 may not all correspond to the q principal eigenvectors of Σ_s . Nonetheless S-PCA has an edge over standard PCA when colored noise corrupts the training data. If the observation noise is white, the eigenspaces of Σ_x and Σ_s coincide and the standard PCA will return the q principal eigenvectors of Σ_s . However, if \mathbf{w}_t is colored the q principal eigenvectors of Σ_x , namely $\{\mathbf{u}_{x,\rho}\}_{\rho=1}^q$, will be different from $\{\mathbf{u}_{s,\rho}\}_{\rho=1}^q$ and may not be sparse. Actually, in *standard PCA* (cf. $\lambda_\rho = 0$) the magnitude of $\|\mathbf{C}_{e,\rho}^T(\bar{\mathcal{S}}_{i_\rho})\|_1$ depends on Σ_w and cannot be made arbitrarily small. Thus, the magnitude between the entries of $\mathbf{C}_{e,\rho}^T$ with indices in $\bar{\mathcal{S}}_{i_\rho}$ relative to those with indices in \mathcal{S}_{i_ρ} cannot be controlled, for a given noise covariance matrix. This prevents one from discerning zero from nonzero entries in $\mathbf{C}_{e,\rho}^T$, meaning that standard PCA cannot guarantee recovery even of a subset of the support of \mathbf{u}_{s,i_ρ} .

On the other hand, Prop. 1 states that S-PCA is capable of identifying a subset of (or all) the support index set of $\{\mathbf{u}_{s,i_\rho}\}_{\rho=1}^q$. S-PCA is more resilient to colored noise than standard PCA because it exploits the sparsity present in the eigenvectors of Σ_s . Intuitively, the ℓ_1 regularization terms act as prior information facilitating emergence of the $\mathbf{U}_{s,r}$ zero entries in \mathbf{C}_e and \mathbf{B}_e as long as the noise variance is not high. Although $\Delta(\Sigma_s)$ has not been explicitly quantified, the upshot of Prop. 1 is that S-PCA is expected to estimate better the columns of $\mathbf{U}_{s,r}$ when compared to standard PCA under comparable noise power. Numerical tests will demonstrate that S-PCA achieves a smaller reconstruction MSE even in the presence of colored noise.

B. Oracle Properties

Turning now attention to the noiseless scenario ($\Sigma_x = \Sigma_s$), S-PCA is expected to perform satisfactorily as long as it estimates well the q principal eigenvectors of Σ_s . Reliable estimators of the clairvoyant matrices $\mathbf{B}_o = \mathbf{C}_o^T = \mathbf{U}_{s,q}^T$ can be obtained when a growing number of training vectors ensures that: i) the probability of identifying the zero entries of the eigenvectors approaches one; and also ii) the estimators of the non-zero entries of $\mathbf{U}_{s,q}$ satisfy a weak form of consistency [35]. Scaling rules for the λ_ρ 's will be derived to ensure that the S-PCA estimates $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ satisfy these so-termed oracle properties. The forthcoming results will be established for the BCD-SPCA scheme (of Sec. III-B), but similar arguments

can be used to prove related claims for ECD-SPCA.

To this end, consider a weighted ℓ_1 -norm in (5), where the sparsity-controlling coefficient multiplying $|\mathbf{B}(j, \rho)|$ and $|\mathbf{C}(\rho, j)|$, namely $\lambda_{\rho,n}$, is replaced by the product $\hat{w}_{j,\rho,n}\lambda_{\rho,n}$. Note the dependence of $\lambda_{\rho,n}$ on n , while the w 's are set equal to $\hat{w}_{j,\rho,n} := |\hat{\mathbf{U}}_{s,n}(j, \rho)|^{-\gamma}$, with $\gamma > 0$ and $\hat{\mathbf{U}}_s$ denoting the estimate of \mathbf{U}_s obtained via standard PCA. If $\mathbf{U}_s(j, \rho)$ is zero, then for n sufficiently large the estimate $\hat{\mathbf{U}}_{s,n}(j, \rho)$ will have a small magnitude. This means large weight $\lambda_{\rho,n}\hat{w}_{j,\rho,n}$ and thus strongly encouraged sparsity in the corresponding estimates $\hat{\mathbf{B}}(j, \rho)$ and $\hat{\mathbf{C}}(\rho, j)$. The oracle properties for $\hat{\mathbf{B}}_{\tau,n}$ and $\hat{\mathbf{C}}_{\tau,n}$ are stated next and proved in Appendix D.

Proposition 2 *Let $\hat{\mathbf{C}}_{\tau,n}$ in (8) be an asymptotically normal estimator of $\mathbf{C}_o = \mathbf{U}_{s,q}^T$; that is,*

$$\hat{\mathbf{C}}_{\tau,n} = \mathbf{U}_{s,q}^T + \sqrt{n^{-1}}\mathbf{E}_{\tau,n}^c \quad (22)$$

where the j th column of $\mathbf{E}_{\tau,n}^c$ converges in distribution, as $n \rightarrow \infty$, to $\mathcal{N}(\mathbf{0}, \Sigma_{E_c,j})$, i.e., a zero-mean Gaussian with covariance $\Sigma_{E_c,j}$. If the sparsity-controlling coefficients are chosen so that

$$\lim_{n \rightarrow \infty} \frac{\lambda_{\rho,n}}{\sqrt{n}} = 0, \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\lambda_{\rho,n}}{n^{(1-\gamma)/2}} = \infty \quad (23)$$

then it holds under (a1) that (8) yields an asymptotically normal estimator of $\mathbf{B}_o = \mathbf{U}_{s,q}$; that is,

$$\hat{\mathbf{B}}_{\tau,n} = \mathbf{U}_{s,q} + \sqrt{n^{-1}}\mathbf{E}_{\tau,n}^b \quad (24)$$

where $\text{vec}(\mathbf{E}_{\tau,n}^b)(\mathcal{S}_o) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(\mathbf{0}, [\Sigma_{E_b}]_{\mathcal{S}_o})$ (convergence in distribution); $[\Sigma_{E_b}]_{\mathcal{S}_o}$ denotes the submatrix formed by the rows and columns, with indices in $\mathcal{S}_o := \mathcal{S}(\mathbf{U}_{s,q}) = \mathcal{S}(\text{vec}(\mathbf{U}_{s,q}))$, of the error covariance Σ_{E_b} of $\text{vec}(\mathbf{E}_{\tau,n}^b)$ obtained when $\hat{\mathbf{B}}_{\tau,n}$ is evaluated by standard PCA. It follows that

$$\lim_{n \rightarrow \infty} \Pr[\mathcal{S}(\hat{\mathbf{B}}_{\tau,n}) = \mathcal{S}(\mathbf{U}_{s,q})] = 1. \quad (25)$$

Following similar arguments as in Prop. 2, it is possible to establish the following corollary.

Corollary 1 *If $\hat{\mathbf{B}}_{\tau,n}$ is an asymptotically normal estimator of \mathbf{B}_o , and the sparsity-controlling coefficients are selected as in Prop. 2, then the optimal solution of (6) at iteration $\tau + 1$, namely $\hat{\mathbf{C}}_{\tau+1,n}$, is an asymptotically normal estimator of \mathbf{C}_o and $\lim_{n \rightarrow \infty} \Pr[\mathcal{S}(\hat{\mathbf{C}}_{\tau+1,n}) = \mathcal{S}(\mathbf{U}_{s,q}^T)] = 1$.*

Prop. 2 and Corollary 1 show that when the BCD-SPCA (or ECD-SPCA) is initialized properly and the sparsity-controlling coefficients follow the scaling rule in (23), then the iterates $\hat{\mathbf{B}}_{\tau,n}$ and $\hat{\mathbf{C}}_{\tau,n}$ satisfy the oracle properties for any iteration index τ . This is important since it shows that the sparsity-aware estimators $\hat{\mathbf{B}}_{\tau,n}$ and $\hat{\mathbf{C}}_{\tau,n}$ achieve MSE performance which asymptotically is as accurate as that attained by a standard PCA approach for the nonzero entries of $\mathbf{U}_{s,q}$. This holds since the error covariance matrix

of the estimates for the nonzero entries of $\mathbf{U}_{s,q}$, namely $[\boldsymbol{\Sigma}_{E_b}]_{S_o}$, coincides with that corresponding to the standard PCA. The estimator $\hat{\mathbf{B}}_{0,n} = \hat{\mathbf{C}}_{0,n}^T = \hat{\mathbf{U}}_{s,q}$ obtained via standard PCA and used to initialize the BCD-SPCA and ECD-SPCA, is asymptotically normal [5], [19].

Remark 3: The scaling laws in (23) resemble those in [35, Thm. 2] for a linear regression problem. The difference here is that the estimate $\hat{\mathbf{B}}_{\tau-1,n}$ (or $\hat{\mathbf{C}}_{\tau,n}$) is nonlinearly related with $\hat{\mathbf{C}}_{\tau,n}$ ($\hat{\mathbf{B}}_{\tau,n}$ respectively). Thus, establishing Prop. 2 requires extra steps to account for the nonlinear interaction between $\hat{\mathbf{C}}_{\tau,n}$ and $\hat{\mathbf{B}}_{\tau,n}$. In order to show asymptotic normality, the chosen weights $\hat{w}_{j,\rho,n}$ are not that crucial. Actually, the part of the proof in Apdx. D that establishes asymptotic normality is valid also when, e.g., $\hat{w}_{j,\rho,n} = 1$ and $\lim_{n \rightarrow \infty} n^{-1/2} \lambda_{\rho,n} = 0$. However, the proposed weights $\hat{w}_{j,\rho,n}$ are instrumental when proving that the probability of recovering the ground-truth support of $\mathbf{U}_{s,q}$ converges to one as the number of training data grows large.

Although the probability of finding the correct support goes to one asymptotically as $n \rightarrow \infty$, numerical tests indicate that this probability is high even when n and τ are finite. This is not the case for the standard PCA estimator, namely $\hat{\mathbf{U}}_{s,q,n} = \text{q-principal eigenvecs}(n^{-1} \sum_{t=1}^n \mathbf{s}_t \mathbf{s}_t^T)$. Numerical examples will also demonstrate that even for a finite number of training data n , the probability of identifying the correct support is increasing as the coordinate descent iteration index τ increases. Thus, the signal subspace estimates obtained by BCD-SPCA (as well as ECD-SPCA) are capable of yielding the correct support of $\mathbf{U}_{s,q}$ even when n is sufficiently large but finite. Consequently, improved estimates of $\mathbf{U}_{s,q}$ are obtained which explains the lower $J_{\text{rec}}(\mathbf{B}, \mathbf{C})$ attained by S-PCA relative to PCA.

V. S-PCA BASED TRANSFORM CODING

Up to this point sparsity has been exploited for DR of data vectors with analog-amplitude entries. However, the majority of modern compression systems are digital. This motivates incorporation of sparsity also in the quantization module that follows DR. This two-stage process comprises the transform coding (TC) approach which has been heavily employed in image compression applications due to its affordable computational complexity [16]. However, current TC schemes do not exploit the presence of sparsity that may be present in the covariance domain.

A sparsity-aware TC (SATC) is proposed here to complement the BCD-SPCA (or ECD-SPCA) algorithm during the data transformation step. The basic idea is to simply quantize the DR vectors using a VQ. Given \mathbf{x} , the DR matrix $\hat{\mathbf{C}}$ obtained by S-PCA is employed during the transformation step to produce the DR vector $\mathbf{y} = \hat{\mathbf{C}}\mathbf{x}$. Then, VQ is employed to produce at the output of the encoder a vector of quantized entries $\hat{\boldsymbol{\eta}}_Q = Q[\mathbf{y}] \in \mathcal{C}_Q^{q \times 1}$, where $\mathcal{C}_Q := \{\hat{\boldsymbol{\eta}}_1, \dots, \hat{\boldsymbol{\eta}}_L\}$ is the quantizer codebook with cardinality $L = 2^R$,

where R denotes the number of bits used to quantize \mathbf{y} . The VQ will be designed numerically using the Max-Lloyd algorithm, as detailed in e.g. [13], which uses $\{\mathbf{y}_t = \hat{\mathbf{C}}\mathbf{x}_t\}_{t=1}^n$ to determine the quantization cells $\{\mathcal{R}_{\hat{\boldsymbol{\eta}}_l}\}_{l=1}^L$, and their corresponding centroids, a.k.a. codewords $\{\hat{\boldsymbol{\eta}}_l\}_{l=1}^L$.

During decoding, the standard process in typical TC schemes [13], [16] is to multiply $\hat{\boldsymbol{\eta}}_Q = Q[\hat{\mathbf{C}}\mathbf{x}]$ with the matrix $(\hat{\mathbf{C}}^T \hat{\mathbf{C}})^\dagger \hat{\mathbf{C}}^T$, and form the estimate $\hat{\mathbf{s}} = (\hat{\mathbf{C}}^T \hat{\mathbf{C}})^\dagger \hat{\mathbf{C}}^T \hat{\boldsymbol{\eta}}_Q$. This estimate minimizes the Euclidean distance $\|\hat{\boldsymbol{\eta}}_Q - \hat{\mathbf{C}}\mathbf{u}\|$ wrt \mathbf{u} . Note that the reconstruction stage of SATC is also used in the DR setting considered in Sections II and III, except that $\hat{\boldsymbol{\eta}}_Q$ is replaced with the vector $\hat{\mathbf{C}}\mathbf{x}$ whose entries are analog. The reason behind using only $\hat{\mathbf{C}}$, and not $\hat{\mathbf{B}}$, is the penalty term $\|\mathbf{B} - \mathbf{C}^T\|_F^2$ which ensures that $\hat{\mathbf{C}}$ and $\hat{\mathbf{B}}^T$ will be close in the ℓ_2 -error sense. Certainly, $\hat{\mathbf{B}}$ could have been used instead, but such a change would not alter noticeably the reconstruction performance. Simulations will demonstrate that the sparsity-inducing mechanisms in the DR step assist SATC to achieve improved MSE reconstruction performance when compared to related sparsity-agnostic TCs.

VI. SIMULATED TESTS

Here the reconstruction performance of ECD-SPCA is studied and compared with the one achieved by standard PCA, as well as sparsity-aware alternatives that were modified to fit the dimensionality reduction setting. The different approaches are compared both in the noiseless and noisy scenarios. Simulation tests are also performed to corroborate the oracle properties established in Sec. IV-B. The SATC is compared with conventional TCs in terms of reconstruction MSE using synthetic data first. Then, SATC is tested in an image compression and denoising application using images from [1].

A. Synthetic Examples

The reconstruction MSE $J_{\text{rec}}(\mathbf{B}_m, \mathbf{C}_m)$ is measured for matrices \mathbf{B}_m and \mathbf{C}_m obtained via: i) ECD-SPCA; ii) the true signal subspace, i.e., $\mathbf{B}_o = \mathbf{C}_o^T = \mathbf{U}_{s,q}$; iii) a ‘sample’-based PCA approach where $\hat{\mathbf{U}}_{x,q}$ is used; iv) a genie-aided PCA which relies on (iii) but also knows where the zero entries of $\mathbf{U}_{s,q}$ are located; v) the sparse PCA approach in [36] abbreviated as ZSPCA; vi) the scheme in [33] abbreviated as SPC; and vii) the algorithm of [9], which is abbreviated as DSPCA. With $p = 14$ and $n = 50$, the MSEs throughout the section are averaged over 200 Monte Carlo runs using a data set that is different from the training set \mathbf{X} . In the noiseless case, $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_s$ is constructed to be a permuted block diagonal matrix with $r = 8$, while 80% of the entries of $\mathbf{U}_{s,r}$ are zero. The sparsity-controlling coefficients multiplying $|\mathbf{B}(j, \rho)|$ and $|\mathbf{C}(\rho, j)|$ are set equal to $\lambda_{\rho,n} |\hat{\mathbf{U}}_x(i, j)|^{-\gamma}$, with $\gamma = 1$ and $\lambda_{\rho,n} \sim n^{0.3}$. Fig. 3 (left) depicts $J_{\text{rec}}(\cdot)$ versus q . The sparsity coefficients in the sparsity-aware approaches are selected from a search grid

to achieve the smallest possible reconstruction MSE. Clearly ECD-SPCA exploits the sparsity present in Σ_s and achieves a smaller reconstruction MSE that is very close to the genie-aided approach. Note that in Fig. 3 (left) there are seven curves. The curve corresponding to sample-based PCA almost overlaps with the one corresponding to SPC. It is also observed that the more sparse the eigenvectors in $\mathbf{U}_{s,r}$ are, the more orthogonal are the ECD-SPCA estimates $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$. This suggests that the ℓ_1 regularization terms in S-PCA induce approximate orthogonality in the corresponding estimates, as long as the underlying eigenvectors forming $\mathbf{U}_{s,r}$ are sufficiently sparse.

Fig. 3 (right) the reconstruction MSE is plotted as a function of the observation SNR, namely $\text{SNR}_{\text{obs}} := 10 \log_{10}[\text{tr}(\Sigma_s)/\text{tr}(\Sigma_w)]$. The colored noise covariance matrix is factored as $\Sigma_w = \mathbf{M}_w \mathbf{M}_w^T$, where \mathbf{M}_w is randomly generated matrix with Gaussian i.i.d. entries. The ECD-SPCA scheme is compared with the sparsity-agnostic standard PCA approach. With $r = q = 3$ and $p = 14$, Σ_s is constructed to be a permuted block diagonal matrix such that 70% of the entries of the eigenmatrix $\mathbf{U}_{s,r}$ are equal to zero. All sparsity coefficients in ECD-SPCA are set equal to $\lambda = 5 * 10^{-3}$. Fig. 3 (right) corroborates that the novel S-PCA can lead to better reconstruction/denoising performance than the standard PCA. The MSE gains are noticeable in the low-to-medium SNR regime. The sparsity imposing mechanisms of ECD-SPCA lead to improved subspace estimates yielding a reconstruction MSE that is close to the one obtained using $\mathbf{U}_{s,q}$. This result corroborates the claims of Prop. 1.

The next three figures validate the S-PCA properties in the noiseless case (see Sec. IV-B). Consider a setting where $p = 14$, $r = 8$, and $q = 2$, and $\Sigma_x = \Sigma_s$ constructed to be a permuted block diagonal matrix such that $\mathbf{U}_{s,r}$ has 70% of its entries equal to zero. The λ 's are selected as in the first paragraph. Fig. 4 (left) displays the signal subspace estimation MSE $E[|||\hat{\mathbf{C}}_{\tau,n}| - |\mathbf{U}_{s,q}^T|||_F^2]$, where the $|\cdot|$ operator is applied entry-wise and is used to eliminate any sign ambiguity present in the rows of $\hat{\mathbf{C}}_{\tau,n}$. As the training data size goes to infinity, the estimation error converges to zero. The convergence speed is similar to the one achieved by standard PCA. Similar conclusions can be deduced for the reconstruction MSE shown in Fig. 4 (right). The reconstruction MSE associated with ECD-SPCA is smaller than the one corresponding to standard PCA. The MSE advantage is larger for a small number of training data in which case standard PCA has trouble locating the zeros of $\mathbf{U}_{s,q}$. These examples corroborate the validity of Prop. 2. Interestingly, multiple coordinate descent iterations ($\tau > 0$) result in smaller estimation and reconstruction MSEs than the one achieved by standard PCA ($\tau = 0$). The MSE gains are noticeable for a small number of training samples. Such gains are expected since ECD-SPCA (or BCD-SPCA) is capable of estimating the true support $\mathcal{S}(\mathbf{U}_{s,q})$ with a positive probability even for a finite n . As shown in Fig. 5 (left), for $\tau > 0$ the probability of finding the true support converges to one as $n \rightarrow \infty$ (cf. Prop. 2). As

τ increases, this probability also increases, while standard PCA ($\tau = 0$) never finds the correct support with a finite number of training data.

Next, the reconstruction MSE of the SATC (Sec.V) is considered and compared with the one achieved by a TC scheme based on standard PCA. The noisy setting used to generate Fig. 3 (right) is considered here with $p = 14$ and $n = 22$. With $r = q = 3$, Σ_s is constructed to be a permuted block diagonal matrix so that 80% of the entries of $\mathbf{U}_{s,r}$ are zero. Data reduction in SATC is performed via ECD-SPCA, while its sparsity controlling coefficients are set as described in the first paragraph of this section. Once the DR matrix $\hat{\mathbf{C}}$ is obtained, the DR training data $\hat{\mathbf{C}}\mathbf{X}$ are used to design the VQ using Max-Lloyd's algorithm. Fig. 5 (right) depicts the reconstruction MSE versus the number of bits used to quantize a single DR vector. Fig. 5 (right) clearly shows that SATC benefits from the presence of sparsity in Σ_s and achieves improved reconstruction performance when compared to the standard TC scheme that relies on PCA. The dashed and solid lines correspond to the reconstruction MSE achieved by ECD-SPCA and $\mathbf{U}_{s,r}$ respectively while no quantization step is present ($R = \infty$).

B. Image compression and denoising

SATC is tested here for compressing and reconstructing images. These images have size 180×256 and they are extracted, as described in Sec. II, from a bigger image of size 2520×2048 in [1]. The images are corrupted with additive zero-mean Gaussian colored noise whose covariance Σ_w is structured as $\Sigma_w = \mathbf{M}\mathbf{M}^T$, where \mathbf{M} contains Gaussian i.i.d. entries. The trace of Σ_w is scaled to fix the SNR at 15dB. Out of a total of 112 generated images, 30 are used for training to determine the DR matrix $\hat{\mathbf{C}}$, and design the VQ. The rest are used as test images to evaluate the reconstruction performance of the following three schemes: i) the SATC; ii) a TC scheme that uses DCT; and iii) a TC scheme which relies on PCA.

The images are split into blocks of dimension 8×8 , and each of the three aforementioned TC schemes is applied to each block. Here $p = 64$ and $\{\mathbf{x}_{t_i}\}_{i=1}^{30}$ is a vectorized representation of an 8×8 sub-block that consists of certain image pixels, while $t_i \in \{1, \dots, 112\}$ denotes the image index. During the operational mode, datum \mathbf{x} corresponds to a noisy sub-block occupying the same row and column indices as the \mathbf{x}_{t_i} 's but belonging to an image that is not in the training set. The signal of interest \mathbf{s} corresponds to the underlying noiseless block we wish to recover. Each noisy datum \mathbf{x} is transformed using either i) the SATC transformation matrix obtained via ECD-SPCA; or ii) the DCT; or iii) the PCA matrix. When the DCT is applied, then DR is performed by keeping the q largest in magnitude entries of the transformed vector. The reduced dimension here is set to $q = 14$. The $q \times 1$ vectors are further quantized by a VQ designed

using the Max-Lloyd algorithm fed with the DR training vectors. At the decoder, the quantized vectors are used to reconstruct \mathbf{s} by: i) using the scheme of Sec. V; ii) multiplying the quantized data with $\hat{\mathbf{U}}_{x,q}$ (PCA); or iii) applying inverse DCT to recover the original block. The sparsity-controlling coefficients in ECD-SPCA are all set equal to $\lambda = 2 * 10^{-2}$.

Fig. 6 shows: a) the original image; b) its noisy version; c) a reconstruction using DCT; d) a reconstruction using a PCA-based TC; and e) the reconstruction returned by SATC. The reconstructed images are obtained after setting the bit rate to $R = 7$. The reconstruction returned by SATC is visually more pleasing than the one obtained by the DCT- and PCA-based TCs. The figure of merit used next is $\text{SNR}_{\text{im}} := 10 \log_{10} \left[\frac{P_{\text{signal}}}{P_{\text{error}}} \right]$, where P_{signal} denotes the power of the noiseless image and P_{error} the power of the noise present in the reconstructed image. Fig. 7 (left) displays SNR_{im} versus the bit-rate of the VQ with SNR set at 17dB. Clearly, SATC achieves higher SNR_{im} values compared to the DCT- and PCA-based TCs. SATC performs better because the ECD-SPCA algorithm used to evaluate the DR matrix $\hat{\mathbf{C}}$ takes advantage of the sparsity present in Σ_s . Similar conclusions can be drawn from Fig. 7 (right) depicting SNR_{im} versus the SNR for $q = 14$ and $R = 7$ bits.

VII. CONCLUDING REMARKS

The present work dealt with compression, reconstruction, and denoising of signal vectors spanned by an orthogonal set of sparse basis vectors that further result a covariance matrix with sparse eigenvectors. Based on noisy training data, a sparsity-aware DR scheme was developed using ℓ_1 -norm regularization to form improved estimates of the signal subspace leading to improved reconstruction performance. Efficient coordinate descent algorithms were developed to minimize the associated non-convex cost. The proposed schemes were guaranteed to converge at least to a stationary point of the cost.

Interesting analytical properties were established for the novel signal subspace estimator showing that even when the noise covariance matrix is unknown, a sufficiently large signal-to-noise ratio ensures that the proposed estimators identify (at least a subset) of the unknown support of the signal covariance eigenvectors. These results advocate that sparsity-aware compression performs well especially when a limited number of training data is available. Asymptotic normality is also established for the sparsity-aware subspace estimators, while it is shown that the probability of these estimates identifying the true signal subspace support approaches one as the number of training data grows large. Appropriate scaling laws for the sparsity-controlling coefficients were derived to satisfy the aforementioned properties.

Finally, the novel S-PCA approach was combined with vector quantization to form a sparsity-aware transform codec (SATC) that was demonstrated to outperform existing sparsity-agnostic approaches. Simulations

using both synthetic data and images corroborated the analytical findings and validated the effectiveness of the proposed schemes. Work is underway to extend the proposed framework to settings involving compression of nonstationary signals, and processes with memory.

ACKNOWLEDGMENT

The authors would like to thank Prof. N. Sidiropoulos of the Technical University of Crete, Greece, for his valuable input and suggestions on the themes of this paper.

APPENDIX

A. *Proof of Lemma 1*: The minimization problem in (12) can be equivalently expressed as

$$\hat{c} = \arg \min_c \|\chi - c\mathbf{h}\|_2^2 + \mu(c - \hat{b})^2 + \lambda t, \quad \text{s. to } -t \leq c \leq t \quad (26)$$

and the derivative of its Lagrangian function involving multipliers ν_1 and ν_2 is given by

$$\nabla_c \mathcal{L}(c, \nu_1, \nu_2) = 2c\mathbf{h}^T \mathbf{h} - 2c\mathbf{h}^T \chi + 2c\mu - 2\mu\hat{b} + \nu_1 - \nu_2.$$

After using the KKT necessary optimality conditions [3, pg. 316], it can be readily deduced that the optimal solution of (12) is given by the second equation in Lemma 1. \square

B. *Proof of convergence of ECD-SPCA*

Let $f(\{\mathbf{B}(j, \rho), \mathbf{C}(\rho, j)\}_{j=1, \rho=1}^{p, q})$ denote the S-PCA cost given in (5), defined over $\mathbb{R}^{2pq \times 1}$; and $f_0(\{\mathbf{B}(j, \rho), \mathbf{C}(\rho, j)\}_{j=1, \rho=1}^{p, q}) := n^{-1} \|\mathbf{X} - \mathbf{B}\mathbf{C}\mathbf{X}\|_F^2 + \mu \|\mathbf{B} - \mathbf{C}^T\|_F^2$. Next, consider the level set

$$\mathcal{F}^0 := \{\{\mathbf{B}(j, \rho), \mathbf{C}(\rho, j)\}_{j=1, \rho=1}^{p, q} : f(\{\mathbf{B}(j, \rho), \mathbf{C}(\rho, j)\}_{j=1, \rho=1}^{p, q}) \leq f(\hat{\mathbf{B}}_0, \hat{\mathbf{C}}_0)\} \quad (27)$$

where $\hat{\mathbf{B}}_0 = \hat{\mathbf{C}}_0^T = \hat{\mathbf{U}}_{s, q}$ correspond to the matrices used to initialize ECD-SPCA obtained via standard PCA. If $\hat{\mathbf{B}}_0$ and $\hat{\mathbf{C}}_0$ have finite ℓ_1 -norms, the set \mathcal{F}^0 is closed and bounded (compact). The latter property can be deduced from (5) and (27), which ensure that matrices \mathbf{B} and \mathbf{C} in \mathcal{F}^0 satisfy $\sum_{\rho, j} \lambda_\rho (|\mathbf{B}(j, \rho)| + |\mathbf{C}(\rho, j)|) \leq f(\hat{\mathbf{B}}_0, \hat{\mathbf{C}}_0)$. Moreover, $f(\hat{\mathbf{B}}_0, \hat{\mathbf{C}}_0)$ is finite when $\hat{\mathbf{B}}_0$ and $\hat{\mathbf{C}}_0$ have finite norms. This is true when the training data \mathbf{X} contain finite entries. Thus, \mathcal{F}^0 is a compact set. Further, the cost function $f(\cdot)$ is continuous on \mathcal{F}^0 .

From (11) and (14) it follows readily that the minimization problems solved to obtain $\hat{\mathbf{C}}_\tau(j, \rho)$ and $\hat{\mathbf{B}}_\tau(\rho, j)$, respectively, are strictly convex. Thus, minimizing $f(\cdot)$ with respect to an entry of \mathbf{C} or \mathbf{B} yields a unique minimizer, namely $\hat{\mathbf{C}}_\tau(\rho, j)$, or $\hat{\mathbf{B}}_\tau(j, \rho)$. Finally, $f(\cdot)$ satisfies the regularization conditions outlined in [30, (A1)]. Specifically, the domain of $f_0(\cdot)$ is formed by matrices whose entries satisfy $\mathbf{B}(j, \rho) \in (-\infty, +\infty)$ and $\mathbf{C}(\rho, j) \in (-\infty, +\infty)$ for $j = 1, \dots, p$ and $\rho = 1, \dots, q$. Thus, $\text{domain}(f_0) =$

$(-\infty, \infty)^{2qp}$ is an *open set*. Moreover, $f_0(\cdot)$ is Gâteaux differentiable over $\text{domain}(f_0)$. Specifically, the Gâteaux derivative of $f_0(\cdot)$ is defined as

$$f'_0(\mathbf{B}, \mathbf{C}; \Delta_B, \Delta_C) := \lim_{h \rightarrow 0} \frac{f_0(\mathbf{B} + h\Delta_B, \mathbf{C} + h\Delta_C) - f_0(\mathbf{B}, \mathbf{C})}{h}.$$

Applying simple algebraic manipulations it follows readily that the Gâteaux derivative exists for all $\Delta_B, \Delta_C \in \text{domain}(f_0)$, and is equal to $2\mu \text{tr}[(\mathbf{B} - \mathbf{C}^T)(\Delta_B - \Delta_C^T)^T] - 2\text{tr}[(\mathbf{X} - \mathbf{B}\mathbf{C}\mathbf{X})(\Delta_B\mathbf{C}\mathbf{X} + \mathbf{B}\Delta_C\mathbf{X})^T]$. Then, convergence of the ECD-SPCA iterates to a stationary point of the S-PCA cost $f(\cdot)$ is readily established using [30, Thm. 4.1 (c)]. \square

C. Proof of Proposition 1: It can be shown by contradiction that for every $\epsilon > 0$ there exists a μ_ϵ such that for any $\mu > \mu_\epsilon$ it holds that $\|\mathbf{B}_e - \mathbf{C}_e^T\|_1 \leq \epsilon/2$. Given that $\Sigma_x = \mathbf{U}_x \mathbf{D}_x \mathbf{U}_x^T$, and since $\|\mathbf{B}_e - \mathbf{C}_e^T\|_1 < \epsilon/2$ for $\mu > \mu_\epsilon$, the minimization problem in (19) can be equivalently rewritten as

$$\mathbf{C}_e(\epsilon) \in \arg \min_{\mathbf{C}} \text{tr} \left(\mathbf{D}_x^{1/2} \mathbf{U}_x^T (\mathbf{I}_q - 2\mathbf{C}^T \mathbf{C} + \mathbf{C}^T \mathbf{C} \mathbf{C}^T \mathbf{C}) \mathbf{U}_x \mathbf{D}_x^{1/2} \right) + \sum_{\rho=1}^q 2\lambda \|\mathbf{C}_{\rho:}^T\|_1 + \phi(\mathbf{C}, \epsilon, \mu) \quad (28)$$

where $\phi(\mathbf{C}, \epsilon, \mu)$ is a continuous function of \mathbf{C} and ϵ , while $\phi(\mathbf{C}, 0, \mu) = 0$.

Let us now consider how the support of each of the rows of \mathbf{C}_e is related to the support of the principal eigenvectors $\{\mathbf{u}_{s,\rho}\}_{\rho=1}^r$. To this end, remove $\phi(\mathbf{C}, \epsilon, \mu)$ from (28) and consider the minimization problem

$$\check{\mathbf{C}}_e \in \arg \min_{\mathbf{C}} \text{tr}[(\mathbf{I} - \mathbf{C}^T \mathbf{C}) \Sigma_x (\mathbf{I} - \mathbf{C}^T \mathbf{C})^T] + \sum_{\rho=1}^q 2\lambda \|\mathbf{C}_{\rho:}^T\|_1. \quad (29)$$

Since the cost in (28) is continuous, one recognizes after applying a continuity argument [11, pg. 15], that for any $\delta > 0$ a sufficiently large μ_δ can be found such that for any $\mu > \max(\mu_\delta, \mu_\epsilon)$ there exists an optimal solution $\mathbf{B}_e, \mathbf{C}_e$ in (28), as well as an optimal solution $\check{\mathbf{C}}_e$ in (29) such that $\|\mathbf{C}_e - \check{\mathbf{C}}_e\|_1 \leq \delta/2$ and $\|\mathbf{B}_e - \check{\mathbf{C}}_e^T\|_1 \leq \|\mathbf{B}_e - \mathbf{C}_e^T\|_1 + \|\mathbf{C}_e - \check{\mathbf{C}}_e\|_1 \leq \delta$ (details are omitted due to space limitations). As the optimal solutions of (28) and (29) can be arbitrarily close, one considers the simpler of the two in (29).

Given that $\Sigma_x = \Sigma_s + \Sigma_w = \mathbf{U}_{s,r} \mathbf{D}_{s,r} \mathbf{U}_{s,r}^T + \mathbf{U}_w \mathbf{D}_w \mathbf{U}_w^T$, the minimization in (29) can be rewritten as

$$\tilde{\mathbf{C}}_e \in \arg \min_{\tilde{\mathbf{C}}} \text{tr} \left(\tilde{\mathbf{C}} (\Sigma_{s,P} + \Sigma_{w,P}) \tilde{\mathbf{C}}^T (\tilde{\mathbf{C}} \tilde{\mathbf{C}}^T - 2\mathbf{I}_q) \right) + 2\lambda \sum_{\rho=1}^q \|\tilde{\mathbf{C}}_{\rho:}^T\|_1 \quad (30)$$

where $\tilde{\mathbf{C}} = \mathbf{C}\mathbf{P}^T$, $\Sigma_{s,P} := \mathbf{P}\Sigma_s\mathbf{P}^T$, $\Sigma_{w,P} := \mathbf{P}\Sigma_w\mathbf{P}^T$ while \mathbf{P} is a permutation matrix constructed so that $\Sigma_{s,P}$ is block diagonal, and $\tilde{\mathbf{C}}_e = \check{\mathbf{C}}_e\mathbf{P}^T$ denotes one of the optimal solutions of (30). Since the ℓ_1 -norm is permutation invariant, it holds that $\|\mathbf{C}_{\rho:}^T\|_1 = \|\tilde{\mathbf{C}}_{\rho:}^T\|_1$.

The minimization problem in (30) can be equivalently written as

$$\tilde{\mathbf{C}}_e \in \arg \min_{\tilde{\mathbf{C}}} \text{tr} \left(\tilde{\mathbf{C}} \Sigma_{x,P} \tilde{\mathbf{C}}^T (\tilde{\mathbf{C}} \tilde{\mathbf{C}}^T - 2\mathbf{I}_q) \right) + 2\lambda \sum_{\rho=1}^q \sum_{j=1}^p \mathbf{T}(\rho, j), \text{ s. to } |\tilde{\mathbf{C}}(\rho, j)| \leq \mathbf{T}(\rho, j). \quad (31)$$

Let \mathbf{T} denote the $q \times p$ matrix whose (ρ, j) th entry is equal to $\mathbf{T}(\rho, j)$. Then, the Lagrangian of (31) is

$$\mathcal{L}(\tilde{\mathbf{C}}, \mathbf{T}, \mathbf{L}_1, \mathbf{L}_2) = \text{tr}(\tilde{\mathbf{C}}\Sigma_{x,P}\tilde{\mathbf{C}}^T(\tilde{\mathbf{C}}\tilde{\mathbf{C}}^T - 2\mathbf{I}_q)) + 2\lambda\mathbf{1}_{q \times 1}^T \mathbf{T}\mathbf{1}_{p \times 1} + \text{tr}(\mathbf{L}_1^T(\tilde{\mathbf{C}} - \mathbf{T})) + \text{tr}(\mathbf{L}_2^T(-\tilde{\mathbf{C}} - \mathbf{T})), \quad (32)$$

where $\mathbf{L}_1, \mathbf{L}_2 \in \mathbb{R}^{q \times p}$ and their (ρ, j) th entry contains the Lagrange multiplier associated with the constraints $\tilde{\mathbf{C}}(\rho, j) \leq \mathbf{T}(\rho, j)$ and $-\tilde{\mathbf{C}}(\rho, j) \leq \mathbf{T}(\rho, j)$, respectively. The first-order optimality conditions imply that the gradient of $\mathcal{L}(\cdot)$ wrt $\tilde{\mathbf{C}}$ should be equal to zero when evaluated at $\tilde{\mathbf{C}}_e$, i.e.,

$$2\tilde{\mathbf{C}}_e\tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e\Sigma_{x,P} + 2\tilde{\mathbf{C}}_e\Sigma_{x,P}\tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e - 4\tilde{\mathbf{C}}_e\Sigma_{x,P} + \mathbf{L}_1 - \mathbf{L}_2 = \mathbf{0}_{q \times p}. \quad (33)$$

Similarly, the gradient of $\mathcal{L}(\cdot)$ wrt \mathbf{T} should be equal to zero at the optimum solution, \mathbf{T}^* , which leads to

$$\mathbf{L}_1^* + \mathbf{L}_2^* = 2\lambda\mathbf{1}_{q \times 1}\mathbf{1}_{p \times 1}^T. \quad (34)$$

Moreover, the optimal multipliers should be nonnegative, i.e., $\mathbf{L}_1^*(\rho, j) \geq 0$ and $\mathbf{L}_2^*(\rho, j) \geq 0$ for $j, \rho = 1, \dots, q$, while the complementary slackness conditions give that $\mathbf{L}_1^*(\rho, j)(\tilde{\mathbf{C}}_e(\rho, j) - \mathbf{T}^*(\rho, j)) = 0$ and $\mathbf{L}_2^*(\rho, j)(-\tilde{\mathbf{C}}_e(\rho, j) - \mathbf{T}^*(\rho, j)) = 0$ (see e.g., [3, pg. 316]).

Let $\mathbf{e}_\rho \in \mathbb{R}^{q \times 1}$ denote the canonical vector which has a single nonzero entry equal to one at the ρ -th position. After multiplying the left hand side (lhs) of (33) from the left with \mathbf{e}_ρ^T and from the right with $\tilde{\mathbf{C}}_{e,\rho}$ we obtain

$$\begin{aligned} 2\tilde{\mathbf{C}}_{e,\rho}^T\tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e\Sigma_{x,P}\tilde{\mathbf{C}}_{e,\rho} + 2\tilde{\mathbf{C}}_{e,\rho}^T\Sigma_{x,P}\tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e\tilde{\mathbf{C}}_{e,\rho} - 4\tilde{\mathbf{C}}_{e,\rho}^T\Sigma_{x,P}\tilde{\mathbf{C}}_{e,\rho} \\ + \sum_{j=1}^p(\mathbf{L}_1^*(\rho, j) - \mathbf{L}_2^*(\rho, j))\tilde{\mathbf{C}}_e(\rho, j) = 0. \end{aligned} \quad (35)$$

Note that the last summand in (35) is equal to $2\lambda\|\tilde{\mathbf{C}}_{\rho,\cdot}\|_1$. This follows from the aforementioned slackness conditions. Specifically, if $\tilde{\mathbf{C}}_e(\rho, j) > 0$ then $\tilde{\mathbf{C}}_e(\rho, j) = \mathbf{T}^*(\rho, j) > 0$, which further implies that $\mathbf{L}_2^*(\rho, j) = 0$ and from (34) it follows that $\mathbf{L}_1^*(\rho, j) = 2\lambda$. In the same way if $\tilde{\mathbf{C}}_e(\rho, j) < 0$, then $\tilde{\mathbf{C}}_e(\rho, j) = -\mathbf{T}^*(\rho, j) < 0$ from which it follows that $\mathbf{L}_1^*(\rho, j) = 0$, thus from (34) we conclude that $\mathbf{L}_2^*(\rho, j) = 2\lambda$. Thus, $(\mathbf{L}_1^*(\rho, j) - \mathbf{L}_2^*(\rho, j))\tilde{\mathbf{C}}_e(\rho, j) = 2\lambda|\tilde{\mathbf{C}}_e(\rho, j)|$, and after some algebraic manipulations on (35) it follows

$$\tilde{\mathbf{C}}_{e,\rho}^T\Sigma_{x,P}(\mathbf{I}_{p \times p} - \tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e)\tilde{\mathbf{C}}_{e,\rho} = 0.5\lambda\|\tilde{\mathbf{C}}_{e,\rho}\|_1, \quad \rho = 1, \dots, q. \quad (36)$$

Summing the q different equalities in (36) we obtain

$$\text{tr}(\Sigma_{x,P}(\mathbf{I}_{p \times p} - \tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e)\tilde{\mathbf{C}}_e^T\tilde{\mathbf{C}}_e) = 0.5\lambda \sum_{\rho=1}^q \|\tilde{\mathbf{C}}_{e,\rho}\|_1 \quad (37)$$

Equality (37) can be used to reformulate the cost in (30) without affecting the optimal solution. Specifically, the cost in (30) can be rewritten as $\text{tr}(\tilde{\mathbf{C}}\Sigma_{x,P}\tilde{\mathbf{C}}^T\tilde{\mathbf{C}}\tilde{\mathbf{C}}^T) - 2\text{tr}(\tilde{\mathbf{C}}\Sigma_{x,P}\tilde{\mathbf{C}}^T) + 2\lambda \sum_{\rho=1}^q \|\tilde{\mathbf{C}}_{\rho,\cdot}\|_1 =$

$-\text{tr}(\tilde{\mathbf{C}}\Sigma_{x,P}\tilde{\mathbf{C}}^T) + 1.5\lambda \sum_{\rho=1}^q \|\tilde{\mathbf{C}}_{\rho,:}^T\|_1$. Using the latter cost expression and expanding the lhs of the q different equality constraints in (36) the minimization problem in (30), is equivalent to

$$\begin{aligned} \tilde{\mathbf{C}}_e \in \arg \min & - \sum_{\rho=1}^q \tilde{\mathbf{C}}_{\rho,:}^T \Sigma_{x,P} \tilde{\mathbf{C}}_{\rho,:} + 1.5\lambda \sum_{\rho=1}^q \|\tilde{\mathbf{C}}_{\rho,:}^T\|_1 \\ \text{s. to } & \left(1 - \|\tilde{\mathbf{C}}_{\rho,:}\|_2^2\right) \tilde{\mathbf{C}}_{\rho,:}^T \Sigma_{x,P} \tilde{\mathbf{C}}_{\rho,:} - \sum_{j=1, j \neq \rho}^q \left(\tilde{\mathbf{C}}_{j,:}^T \tilde{\mathbf{C}}_{\rho,:}\right) \left(\tilde{\mathbf{C}}_{\rho,:}^T \Sigma_{x,P} \tilde{\mathbf{C}}_{j,:}\right) = 0.5\lambda \|\tilde{\mathbf{C}}_{\rho,:}\|_1, \quad \rho = 1, \dots, q. \end{aligned} \quad (38)$$

Each one of the summands of the sum in the lhs of the equality constraints in (38) can be rewritten as

$$\tilde{\mathbf{C}}_{j,:}^T \tilde{\mathbf{C}}_{\rho,:} \tilde{\mathbf{C}}_{\rho,:}^T \Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} = \tilde{\mathbf{C}}_{\rho,:}^T \left(\Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} \tilde{\mathbf{C}}_{j,:}^T \right) \tilde{\mathbf{C}}_{\rho,:} \geq 0, \quad j, \rho = 1, \dots, q \text{ and } \rho \neq j. \quad (39)$$

Notice that the quantity in (39) is nonnegative since $\text{rank}(\Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} \tilde{\mathbf{C}}_{j,:}^T) = 1$, while the single nonzero eigenvalue of $\Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} \tilde{\mathbf{C}}_{j,:}^T$ is $d_{\max}(\Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} \tilde{\mathbf{C}}_{j,:}^T) = \text{tr}(\Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} \tilde{\mathbf{C}}_{j,:}^T) = \tilde{\mathbf{C}}_{j,:}^T \Sigma_{x,P} \tilde{\mathbf{C}}_{j,:} > 0$ for $\tilde{\mathbf{C}}_{j,:} \neq 0$. From the constraints in (38) and (39), it follows that $\|\tilde{\mathbf{C}}_{\rho,:}\|_2 \leq 1$, otherwise $\|\tilde{\mathbf{C}}_{\rho,:}\|_1$ would be negative resulting a contradiction.

For the time being let us ignore the noise covariance matrix $\Sigma_{w,P}$ by setting it to zero, thus $\Sigma_{x,P} = \Sigma_{s,P}$. For the selected sparsity-controlling coefficient λ in (30) assume that the optimal solution has $\|\tilde{\mathbf{C}}_{e,\rho,:}^T\|_0 = l_\rho$ and $\|\tilde{\mathbf{C}}_{e,\rho,:}^T\|_1 = \kappa_\rho$, while $2 \leq l_\rho \leq G$ for $\rho = 1, \dots, q$. This is possible since the ℓ_1 -norm is used in S-PCA. The case $q = 1$ is considered first to demonstrate the main result which is then generalized for $q > 1$. Toward this end, let $\tilde{\mathbf{C}}_{1,:} = \|\tilde{\mathbf{C}}_{1,:}\|_2^2 \cdot \mathbf{u}_{\tilde{e},1}$, where $\|\mathbf{u}_{\tilde{e},1}\|_2 = 1$. Moreover, to simplify notation let $c_1 = \|\tilde{\mathbf{C}}_{1,:}\|_2^2$, and $\gamma_1 = \mathbf{u}_{\tilde{e},1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},1}$; thus $\tilde{\mathbf{C}}_{1,:}^T \Sigma_{s,P} \tilde{\mathbf{C}}_{1,:} = c_1^2 \gamma_1 \geq 0$. Let d_1^* denote the maximum spectral radius among all possible $l_1 \times l_1$ submatrices of $\Sigma_{s,P}$ that are formed after keeping l_1 of its rows and columns with common indices that are determined by the indices of the l_1 nonzero entries in the optimal $\tilde{\mathbf{C}}_{e,1,:} = \|\tilde{\mathbf{C}}_{e,1,:}\|_2^2 \mathbf{u}_{\tilde{e},1}$, and $\mathbf{u}_{\tilde{e},1}$ the optimal selection for $\mathbf{u}_{\tilde{e},1}$. Then, it holds that $\gamma_1 = \mathbf{u}_{\tilde{e},1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},1} \leq d_1^*$ for any unit-vector $\mathbf{u}_{\tilde{e},1}$ for which $\|\mathbf{u}_{\tilde{e},1}\|_0 = l_1$. With this notation in mind, $q = 1$ and $\|\tilde{\mathbf{C}}_{e,1,:}^T\|_1 = \kappa_1$ (38) is equivalent to

$$\min_{c_1, \gamma_1} -c_1^2 \gamma_1, \quad \text{s. to } (1 - c_1^2) c_1^2 \gamma_1 = 0.5\lambda \kappa_1, \quad 0 \leq c_1 \leq \kappa_1' = \min(1, \kappa_1), \quad 0 \leq \gamma_1 \leq d_1^*, \quad (40)$$

where the first inequality constraint in (40) follows from the fact that $c_1 \leq 1$ and $c_1 = \|\tilde{\mathbf{C}}_{1,:}\|_2 < \kappa_1 = \|\tilde{\mathbf{C}}_{1,:}\|_1$. The Lagrangian of (40) is given as

$$\mathcal{L}_1(c_1, \gamma_1, \mathbf{v}) = -c_1^2 \gamma_1 + v_1^a [(1 - c_1^2) c_1^2 \gamma_1 - 0.5\lambda \kappa_1] + v_1^b [c_1 - \kappa_1'] - v_1^c c_1 - v_1^d \gamma_1 + v_1^e (\gamma_1 - d_1^*), \quad (41)$$

where $\mathbf{v} := [v_1^a \ v_1^b \ v_1^c \ v_1^d \ v_1^e]^T$ contains the Lagrange multipliers. After i) differentiating (41) with respect to c_1 and γ_1 ; ii) setting the corresponding derivatives equal to zero; and iii) applying the complementary slackness conditions [see also Karush-Kuhn-Tucker necessary optimality condition in [3, pg. 316]] it

follows that the optimal value $v_1^{e,*}$ of the multiplier v_1^e should be strictly positive. The slackness conditions imply that $v_1^{e,*}(\gamma_1^* - d_1^*) = 0$, then it follows that at the minimum of (40) it holds that $\gamma_1^* = d_1^*$. Now recall that $\gamma_1 = \mathbf{u}_{\tilde{e},1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},1}$, thus γ_1^* is formed when $\mathbf{u}_{\tilde{e},1} = \mathbf{u}_{\tilde{e},e,1}$ (the optimal direction toward which the optimal row $\mathbf{C}_{e,\rho}$ is pointing).

Recall that $d_1^* = \max_{\mathbf{u}_{\tilde{e},1}} \mathbf{u}_{\tilde{e},1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},1}$ subject to $\|\mathbf{u}_{\tilde{e},1}\|_2 = 1$ and $\|\mathbf{u}_{\tilde{e},1}\|_0 = l_1$. Next, we demonstrate that if $\mathbf{u}_{\tilde{e},e,1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},e,1} = d_1^*$, then there exists a column, say the i_1 th in $\tilde{\mathbf{U}}_{s,r}$, with support $\tilde{\mathcal{S}}_{i_1}$ such that $\|\mathbf{u}_{\tilde{e},e,1}^T(\tilde{\mathcal{S}}_{i_1})\|_1 = 0$, while $\|\mathbf{u}_{\tilde{e},e,1}^T(\tilde{\mathcal{S}}_{i_1})\|_1 > 0$ and $\tilde{\mathcal{S}}_{i_1}$ denotes the complement of $\tilde{\mathcal{S}}_{i_1}$. Since $\tilde{\mathbf{C}}_{e,1}^T$ is a scaled version of $\mathbf{u}_{\tilde{e},e,1}^T$, the latter property will further imply that $\|\tilde{\mathbf{C}}_{e,1}^T(\tilde{\mathcal{S}}_{i_1})\|_1 = 0$, while $\|\tilde{\mathbf{C}}_{e,1}^T(\tilde{\mathcal{S}}_{i_1})\|_1 \geq \xi'_1(\lambda)$, where $\xi'_1(\lambda)$ is strictly positive. Equivalently, we will show that $\mathcal{I}_1 := \mathcal{S}(\tilde{\mathbf{C}}_{e,1}) = \mathcal{S}(\mathbf{u}_{\tilde{e},e,1}) \subseteq \mathcal{G}_{k_1}$, where $\mathcal{G}_{k_1} = \tilde{\mathcal{S}}_{i_1}$ corresponds to the index set of the entries of $\mathbf{s}_P = \mathbf{P}\mathbf{s}$ that belong to, say the k_1 th diagonal block of $\Sigma_{s,P}$ and $k_1 \in \{1, \dots, K\}$. To this end, let $\mathbf{u}_{\tilde{e},1}^T = [(\mathbf{u}_{\tilde{e},1}^1)^T, (\mathbf{u}_{\tilde{e},1}^2)^T, \dots, (\mathbf{u}_{\tilde{e},1}^K)^T]$, where each subvector $\mathbf{u}_{\tilde{e},1}^k$ has G entries; and let $\mathcal{I}_{1,k} := \mathcal{S}(\mathbf{u}_{\tilde{e},1}^k)$ with $\sum_{k=1}^K |\mathcal{I}_{1,k}| = l_1$. Then, it follows that

$$\mathbf{u}_{\tilde{e},1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},1} = \sum_{k=1}^K (\mathbf{u}_{\tilde{e},1}^k)^T \Sigma_{s_{\mathcal{G}_k}} \mathbf{u}_{\tilde{e},1}^k \leq \sum_{k=1}^K d_{\max}(\Sigma_{s_{\mathcal{G}_k}}^{l_1}) \|\mathbf{u}_{\tilde{e},1}^k\|_2^2, \quad (42)$$

where $d_{\max}(\Sigma_{s_{\mathcal{G}_k}}^{l_1})$ denotes the spectral radius of the $l_1 \times l_1$ submatrix $\Sigma_{s_{\mathcal{G}_k}}^{l_1}$ formed by the $G \times G$ diagonal block $\Sigma_{s_{\mathcal{G}_k}}$ of $\Sigma_{s,P}$ after keeping l_1 of its rows and columns with common indices. The inequality in (42) follows since each subvector $\mathbf{u}_{\tilde{e},1}^k$ of $\mathbf{u}_{\tilde{e},1}$ can have at most l_1 nonzero entries. If $d_1^{l_1}$ denotes the maximum spectral radius that can be achieved by any $l_1 \times l_1$ submatrix $\Sigma_{s_k}^{l_1}$ that is contained in a diagonal block $\Sigma_{s_{\mathcal{G}_k}}$ of $\Sigma_{s,P}$ for $k = 1, \dots, K$, then from (42) and since $\sum_{k=1}^K \|\mathbf{u}_{\tilde{e},1}^k\|_2^2 = 1$ it holds that $\mathbf{u}_{\tilde{e},1}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},1} \leq d_1^{l_1}$. Thus, it should hold that $d_1^{l_1} = d_1^*$. Then, the max value d_1^* can be attained if and only if the indices of the nonzero entries of $\mathbf{u}_{\tilde{e},e,1}$ satisfy $\mathcal{I}_1 \subseteq \mathcal{G}_{k_1}$ for a $k_1 \in \{1, \dots, K\}$. This further implies that there exists an eigenvector $\tilde{\mathbf{u}}_{s,i_1} := \mathbf{P}\mathbf{u}_{s,i_1}$ with support $\tilde{\mathcal{S}}_{i_1} = \mathcal{G}_{k_1}$, for which $\mathcal{I}_1 \subseteq \tilde{\mathcal{S}}_{i_1}$. Thus, it is deduced that $\mathbf{u}_{\tilde{e},e,1}^T(\tilde{\mathcal{S}}_{i_1}) = 0$ and $\|\mathbf{u}_{\tilde{e},e,1}^T(\tilde{\mathcal{S}}_{i_1})\|_1 \geq \xi'_1(\lambda) > 0$ since the l_1 nonzero entries of $\mathbf{u}_{\tilde{e},e,1}^T$ have indices in $\tilde{\mathcal{S}}_{i_1}$. Positivity of $\xi'_1(\lambda)$ is ensured since $\|\mathbf{u}_{\tilde{e},e,1}\|_2 = 1$ and λ is selected such that $\tilde{\mathbf{C}}_{e,1} = (c_1^*)^2 \mathbf{u}_{\tilde{e},e,\rho} \neq \mathbf{0}$. Since $\tilde{\mathbf{C}}_e = \tilde{\mathbf{C}}_e \mathbf{P}$ in (29) results from permuting the columns of $\tilde{\mathbf{C}}_e$, it follows that $\|\tilde{\mathbf{C}}_{e,1}^T(\tilde{\mathcal{S}}_{i_\rho})\|_1 = 0$ and $\|\tilde{\mathbf{C}}_{e,1}^T(\mathcal{S}_{i_\rho})\|_1 \geq \xi'_1(\lambda) > 0$, where $\mathcal{S}_{i_\rho} = \mathcal{S}(\mathbf{u}_{s,i_\rho})$.

We generalize the previous claim for the case when $q > 1$. As before we reexpress each of the rows of $\tilde{\mathbf{C}}$ as $\tilde{\mathbf{C}}_\rho = \|\tilde{\mathbf{C}}_\rho\|_2 \mathbf{u}_{\tilde{e},\rho}$, with $\|\mathbf{u}_{\tilde{e},\rho}\|_2 = 1$ for $\rho = 1, \dots, q$. No other assumptions are imposed for the direction vectors $\mathbf{u}_{\tilde{e},\rho}$. Further, let $c_\rho = \|\tilde{\mathbf{C}}_\rho\|_2$ and $\gamma_\rho = \mathbf{u}_{\tilde{e},\rho}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},\rho}$, where $\rho = 1, \dots, q$. Moreover, let $\delta_{j,\rho} = (\mathbf{u}_{\tilde{e},j}^T \mathbf{u}_{\tilde{e},\rho})(\mathbf{u}_{\tilde{e},j}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},\rho})$ for $j \neq \rho$ and further notice that $\delta_{j,\rho} = \delta_{\rho,j} \geq 0$ [cf. (39)], while $\delta_{j,\rho} \leq \gamma_\rho$ and $\delta_{j,\rho} \leq \gamma_j$. Also recall that λ has been selected such that $\{\|\tilde{\mathbf{C}}_{e,\rho}\|_1 = \kappa_\rho\}_{\rho=1}^q$

and $\{\|\tilde{\mathbf{C}}_{e,\rho}\|_0 = l_\rho\}_{\rho=1}^q$ where $2 \leq l_\rho \leq G$. Then, the minimization problem in (38) can be equivalently rewritten as

$$\begin{aligned} \min - \sum_{\rho=1}^q c_\rho^2 \gamma_\rho, \quad \text{s. to } (1 - c_\rho^2) c_\rho^2 \gamma_\rho - \sum_{j=1, j \neq \rho}^q c_\rho^2 c_j^2 \delta_{j,\rho} = 0.5 \lambda \kappa_\rho, \quad 0 \leq c_\rho \leq \kappa'_\rho := \min(1, \kappa_\rho) \\ 0 \leq \gamma_\rho \leq d_\rho^*, \quad 0 \leq \delta_{j,\rho} \leq \gamma_\rho, \quad \delta_{j,\rho} \leq \gamma_j, \quad \delta_{j,\rho} = \delta_{\rho,j}, \quad j \neq \rho, \quad j, \rho = 1, \dots, q \end{aligned} \quad (43)$$

where d_ρ^* corresponds to the maximum value that $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{\Sigma}_{s,P} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$ can attain when $\|\tilde{\mathbf{C}}_{e,\rho}\|_0 = l_\rho$, while $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$ are selected such that the constraints in (43) are satisfied. The Lagrangian function of (43) is given as

$$\begin{aligned} \mathcal{L}_2(\{c_\rho, \gamma_\rho\}_{\rho=1}^q, \{\delta_{j,\rho}\}, \mathbf{v}) = & - \sum_{\rho=1}^q c_\rho^2 \gamma_\rho + \sum_{\rho=1}^q v_\rho^a [(1 - c_\rho^2) c_\rho^2 \gamma_\rho - \sum_{j=1, j \neq \rho}^q c_\rho^2 c_j^2 \delta_{j,\rho} - 0.5 \lambda \kappa_\rho] \\ & + \sum_{\rho=1}^q \left[v_\rho^b (c_\rho - \kappa'_\rho) - v_\rho^c c_\rho + v_\rho^d (\gamma_\rho - d_\rho^*) - v_\rho^e \gamma_\rho \right] \\ & + \sum_{\rho=1}^q \sum_{j=1, j \neq \rho}^q \left[v_{j,\rho}^f (\delta_{j,\rho} - \gamma_\rho) + v_{j,\rho}^g (\delta_{j,\rho} - \gamma_j) + v_{j,\rho}^h (\delta_{j,\rho} - \delta_{\rho,j}) - v_{j,\rho}^i \delta_{j,\rho} \right], \end{aligned} \quad (44)$$

where \mathbf{v} is a vector that contains the Lagrange multipliers $v_\rho^a, v_\rho^b, v_\rho^c, v_\rho^d, v_\rho^e, v_{j,\rho}^f, v_{j,\rho}^g, v_{j,\rho}^h$ and $v_{j,\rho}^i$. The KKT conditions are applied next to derive necessary conditions that the optimal solution of (43) should satisfy. This involves i) differentiating (44) wrt c_ρ, γ_ρ and $\delta_{j,\rho}$; ii) setting the corresponding derivatives equal to zero; and iii) applying the complementary slackness conditions for the optimal multipliers \mathbf{v}^* . Then, it follows that at the minimum of (43) it should hold that $\delta_{j,\rho}^* = 0$, and $\gamma_\rho^* = d_\rho^*$ for $j, \rho = 1, \dots, q$ and $j \neq \rho$. From the definition of $\delta_{j,\rho}$ it follows that $\delta_{j,\rho}^*$ is formed using the optimal vectors $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$, i.e., $\delta_{j,\rho}^* = (\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}})(\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{\Sigma}_{s,P} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}})$. Since $\delta_{j,\rho}^* = 0$, it follows that the optimal direction vector $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$ should be selected in (38) such that $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}} = 0$, or $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{\Sigma}_{s,P} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}} = 0$ for $j \neq \rho$, while $\gamma_\rho^* = \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{\Sigma}_{s,P} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$ is equal to the maximum possible value d_ρ^* . Since $\tilde{\mathbf{C}}_{e,\rho} = (c_\rho^*)^2 \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$ it follows that the rows of the optimal matrix $\tilde{\mathbf{C}}_e$ should be selected such that either they are orthogonal $\tilde{\mathbf{C}}_{e,\rho}^T \tilde{\mathbf{C}}_{e,j} = 0$, or $\tilde{\mathbf{C}}_{e,\rho}^T \mathbf{\Sigma}_{s,P} \tilde{\mathbf{C}}_{e,j} = 0$. In summary the direction vector for the ρ th row of the optimal matrix $\tilde{\mathbf{C}}_e$ in (38), namely $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}$, should be selected such that

$$\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}} = \arg \max_{\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{\Sigma}_{s,P} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}, \quad \text{s. to } (\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}})(\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T \mathbf{\Sigma}_{s,P} \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}) = 0, \quad \|\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}\|_2 = 1, \quad \|\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}\|_1 = l_\rho,$$

where $\rho = 1, \dots, q, j \neq \rho$. Using similar reasoning as in the case where $q = 1$ it follows that for every optimal row $\tilde{\mathbf{C}}_{e,\rho}$ there exists $i_\rho \in \{1, \dots, r\}$ such that $\mathcal{I}_\rho = \mathcal{S}(\tilde{\mathbf{C}}_{e,\rho}^T) \subseteq \mathcal{S}(\tilde{\mathbf{u}}_{s,i_\rho})$. Letting $\tilde{\mathcal{S}}_{i_\rho}$ be the complement of $\tilde{\mathcal{S}}_{i_\rho}$, it is deduced that $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T (\tilde{\mathcal{S}}_{i_\rho}) = 0$ and $\|\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T (\tilde{\mathcal{S}}_{i_\rho})\|_1 \geq \xi'_\rho(\lambda) > 0$ since the l_ρ nonzero entries of $\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T$ have indices in $\tilde{\mathcal{S}}_{i_\rho}$. Positivity of $\xi'_\rho(\lambda)$ is ensured since $\|\mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}}^T\|_2 = 1$ and $\tilde{\mathbf{C}}_{e,\rho} = (c_\rho^*)^2 \mathbf{u}_{\tilde{\mathbf{C}}_{e,\rho}} \neq \mathbf{0}$ for the selected λ . Then, it follows readily that $\|\tilde{\mathbf{C}}_{e,\rho} (\tilde{\mathcal{S}}_{i_\rho})\|_1 = 0$ while

$\|\tilde{\mathbf{C}}_{e,\rho}(\tilde{\mathcal{S}}_{i_\rho})\|_1 \geq \xi'_\rho(\lambda) > 0$ for $\rho = 1, \dots, q$. Since $\check{\mathbf{C}}_e = \tilde{\mathbf{C}}_e \mathbf{P}$ in (29) results from permuting the columns of $\tilde{\mathbf{C}}_e$ it follows that $\|\check{\mathbf{C}}_{e,\rho}^T(\tilde{\mathcal{S}}_{i_\rho})\|_1 = 0$ and $\|\check{\mathbf{C}}_{e,\rho}^T(\mathcal{S}_{i_\rho})\|_1 \geq \xi'_\rho(\lambda) > 0$, where $\mathcal{S}_{i_\rho} = \mathcal{S}(\mathbf{u}_{s,i_\rho})$.

The latter property was proved under the assumption that $\Sigma_{w,P} = 0$. Consider now the general case where $\Sigma_{w,P} \neq 0$, thus $\Sigma_{x,P} \neq \Sigma_{s,P}$. An upper bound on the noise variance will be determined that ensures the validity of the earlier claims about $\check{\mathbf{C}}_{e,\rho}^T$ (or $\check{\mathbf{C}}_{e,\rho}^T$) established in the noiseless case. Let $\check{\mathbf{u}}_{\tilde{e},\rho}$ be a direction vector that results a row vector $\check{\mathbf{C}}_{\rho}$ that belongs to the constraint set of (38), while $\|\check{\mathbf{u}}_{\tilde{e},\rho}\|_0 = \|\check{\mathbf{C}}_{\rho}\|_0 = l_\rho$. Further, assume that the support of $\check{\mathbf{u}}_{\tilde{e},\rho}$ is different from the support of the optimal $\mathbf{u}_{\tilde{e},e,\rho}^T$ evaluated when $\Sigma_{w,P} = 0$. One sufficient condition to ensure optimality of $\{\mathbf{u}_{\tilde{e},e,\rho}^T\}_{\rho=1}^q$ in the presence of noise is that

$$\check{\mathbf{u}}_{\tilde{e},\rho}^T \Sigma_{s,P} \check{\mathbf{u}}_{\tilde{e},\rho} + \check{\mathbf{u}}_{\tilde{e},\rho}^T \Sigma_{w,P} \check{\mathbf{u}}_{\tilde{e},\rho} < d_\rho^* = \mathbf{u}_{\tilde{e},e,\rho}^T \Sigma_{s,P} \mathbf{u}_{\tilde{e},e,\rho}, \quad \rho = 1, \dots, q. \quad (45)$$

for any $\check{\mathbf{u}}_{\tilde{e},\rho}$ that results a feasible $\check{\mathbf{C}}_{\rho}$ in (38), while $\|\check{\mathbf{u}}_{\tilde{e},\rho}\|_0 = l_\rho$ and $\mathcal{S}(\check{\mathbf{u}}_{\tilde{e},\rho}) \neq \mathcal{S}(\mathbf{u}_{\tilde{e},e,\rho})$. Given that $\check{\mathbf{u}}_{\tilde{e},\rho}^T \Sigma_{w,P} \check{\mathbf{u}}_{\tilde{e},\rho} \leq d_{\max}(\Sigma_{w,P})$, it follows that (45) will be satisfied when

$$d_{\max}(\Sigma_{w,P}) < d_\rho^* - \check{\mathbf{u}}_{\tilde{e},\rho}^T \Sigma_{s,P} \check{\mathbf{u}}_{\tilde{e},\rho}. \quad (46)$$

Note that $\check{\mathbf{u}}_{\tilde{e},\rho}^T \Sigma_{s,P} \check{\mathbf{u}}_{\tilde{e},\rho} < d_\rho^*$ since $\check{\mathbf{u}}_{\tilde{e},\rho}$ does not have the same support as $\mathbf{u}_{\tilde{e},e,\rho}^T$ that maximizes the problem at the bottom of pg. 28, in which $\Sigma_{w,P} = 0$. Thus, the quantity in the right hand side of (46), denoted as $\Delta(\Sigma_s)$, will be positive.

What remains to establish are the properties stated in Prop. 1 for $\mathbf{C}_{e,\rho}^T$ and $\mathbf{B}_{e,\rho}$ with $\rho = 1, \dots, q$. To this end, recall that for any $\mu > \max(\mu_\delta, \mu_\epsilon)$, and for each $\tilde{\mathbf{C}}_e$, or equivalently $\check{\mathbf{C}}_e$, there exists an optimal solution \mathbf{C}_e and \mathbf{B}_e of (19) for which $\|\mathbf{C}_e - \check{\mathbf{C}}_e\|_1 \leq \delta/2$ and $\|\mathbf{B}_e - \check{\mathbf{C}}_e^T\|_1 \leq \delta$, where $\check{\mathbf{C}}_e = \tilde{\mathbf{C}}_e \mathbf{P}$. Then, $\|\mathbf{C}_{e,\rho}^T - \check{\mathbf{C}}_e\|_1 \leq \delta/2$ for $\rho = 1, \dots, q$. Then, it readily follows that $\|\mathbf{C}_{e,\rho}^T(\tilde{\mathcal{S}}_{i_\rho})\|_1 \leq \delta/2$ since $\check{\mathbf{C}}_{e,\rho}^T(\tilde{\mathcal{S}}_{i_\rho}) = \mathbf{0}^T$. Moreover, $\|\check{\mathbf{C}}_{e,\rho}^T(\mathcal{S}_{i_\rho})\| - \delta/2 \leq \|\mathbf{C}_{e,\rho}^T(\mathcal{S}_{i_\rho})\|_1 \leq \|\check{\mathbf{C}}_{e,\rho}^T(\mathcal{S}_{i_\rho})\| + \delta/2$. Notice that the lower bound $\|\check{\mathbf{C}}_{e,\rho}^T(\mathcal{S}_{i_\rho})\| - \delta/2 \geq \xi'_\rho(\lambda) - \delta/2$ can be made strictly positive by pushing $\delta/2$ arbitrarily close to zero, which is possible by increasing μ . However, $\xi_\rho(\lambda)$ remains strictly positive for the values of λ_ρ considered here, since it does not depend on μ . These properties can also be established for $\mathbf{B}_{e,\rho}$ using similar arguments. \square

D. Proof of Proposition 2: In the noiseless case the training matrix $\mathbf{X}_n = \mathbf{S}_n$ (note the dependence on n) can be written as $\mathbf{S}_n = \mathbf{U}_{s,q} \mathbf{U}_{s,q}^T \mathbf{S}_n + \mathbf{U}_{s,p-q} \mathbf{U}_{s,p-q}^T \mathbf{S}_n$. For notational convenience let $\Gamma_{q,n} = \mathbf{U}_{s,q}^T \mathbf{S}_n$, and $\mathbf{z}_{q,n} = \text{vec}(\mathbf{U}_{s,p-q} \mathbf{U}_{s,p-q}^T \mathbf{S}_n) = ((\mathbf{U}_{s,p-q}^T \mathbf{S}_n)^T \otimes \mathbf{I}_p) \text{vec}(\mathbf{U}_{s,p-q})$. Using vec notation, it holds that $\text{vec}(\mathbf{S}_n) = (\Gamma_{q,n}^T \otimes \mathbf{I}_p) \text{vec}(\mathbf{U}_{s,q}) + \mathbf{z}_{q,n}$. Moreover, let $\mathbf{b} = \text{vec}(\mathbf{B}) = \text{vec}(\mathbf{U}_{s,q}) + \sqrt{n^{-1}} \tilde{\mathbf{b}}$, where $\sqrt{n^{-1}} \tilde{\mathbf{b}}$ quantifies the estimation error present when estimating $\mathbf{U}_{s,q}$ via (8). Using this notation and after applying

some algebraic manipulations the cost in (8) can be reformulated as

$$\begin{aligned}
J_b(\tilde{\mathbf{b}}) := & \|(\mathbf{\Gamma}_{q,n}^T \otimes \mathbf{I}_p) \text{vec}(\mathbf{U}_{s,q}) + \mathbf{z}_{q,n} - ((\hat{\mathbf{C}}_{\tau,n} \mathbf{S}_n)^T \otimes \mathbf{I}_p) \tilde{\mathbf{b}}\|_2^2 + \sum_{\rho=1}^q \lambda_{\rho,n} \sum_{j_\rho=p(\rho-1)+1}^{\rho p} |\mathbf{b}(j_\rho)| \\
& + \mu \|\mathbf{b} - \hat{\mathbf{c}}_{\tau,n}^t\|_2^2 = \| -\sqrt{n^{-1}} [(\hat{\mathbf{C}}_{\tau,n} \mathbf{S}_n)^T \otimes \mathbf{I}_p] \tilde{\mathbf{b}} + \mathbf{z}_{q,n} - \sqrt{n^{-1}} [(\mathbf{E}_{\tau,n}^c \mathbf{S}_n)^T \otimes \mathbf{I}_p] \text{vec}(\mathbf{U}_{s,q}) \|_2^2 + \\
& \sum_{\rho=1}^q \lambda_{\rho,n} \sum_{j_\rho=(\rho-1)p+1}^{\rho p} |\sqrt{n^{-1}} \tilde{\mathbf{b}}(j_\rho) + \text{vec}(\mathbf{U}_{s,q})(j_\rho)| + \mu \|\sqrt{n^{-1}} \tilde{\mathbf{b}} - \sqrt{n^{-1}} \mathbf{e}_{\tau,n}^c\|_2^2
\end{aligned} \quad (47)$$

where the second inequality follows after replacing \mathbf{b} with $\text{vec}(\mathbf{U}_{s,q}) + \sqrt{n^{-1}} \tilde{\mathbf{b}}$ in all three terms in the expression following $J_b(\tilde{\mathbf{b}})$. Moreover, $\hat{\mathbf{c}}_{\tau,n}^t := \text{vec}(\hat{\mathbf{C}}_{\tau,n}^T)$, $\mathbf{e}_{\tau,n}^c := \text{vec}((\mathbf{E}_{\tau,n}^c)^T)$, and $\text{vec}(\mathbf{U}_{s,q})(j)$ denotes the j th element of $\text{vec}(\mathbf{U}_{s,q})$. Recall that the optimal solution of (8) is $\hat{\mathbf{B}}_{\tau,n}$, and let $\check{\mathbf{b}}_n := \sqrt{n}[\text{vec}(\hat{\mathbf{B}}_{\tau,n}) - \text{vec}(\mathbf{U}_{s,q})]$. We will show that the error $\check{\mathbf{b}}_n$ which minimizes (47) and corresponds to the estimate $\hat{\mathbf{B}}_{\tau,n}$ converges to a Gaussian random variable, thus establishing the first result in Prop. 2.

To this end, consider the cost $J_b(\tilde{\mathbf{b}}) - J_b(\mathbf{0})$ which has the same optimal solution as $J_b(\tilde{\mathbf{b}})$, since $J_b(\mathbf{0})$ is a constant. After performing some algebraic manipulations we can readily obtain

$$\begin{aligned}
J_b(\tilde{\mathbf{b}}) - J_b(\mathbf{0}) = & n^{-1} \tilde{\mathbf{b}}^T \left[(\hat{\mathbf{C}}_{\tau,n} \mathbf{S}_n \mathbf{S}_n^T \hat{\mathbf{C}}_{\tau,n}^T + \mu \mathbf{I}_q) \otimes \mathbf{I}_{p \times p} \right] \tilde{\mathbf{b}} - 2n^{-0.5} \tilde{\mathbf{b}}^T \left[(\hat{\mathbf{C}}_{\tau,n} \mathbf{S}_n \mathbf{S}_n^T \mathbf{U}_{s,p-q}) \otimes \mathbf{I}_p \right] \text{vec}(\mathbf{U}_{s,p-q}) \\
& - 2\mu n^{-1} \tilde{\mathbf{b}}^T \mathbf{e}_{\tau,n}^c + 2n^{-1} \tilde{\mathbf{b}}^T \left[(\hat{\mathbf{C}}_{\tau,n} \mathbf{S}_n \mathbf{S}_n^T (\mathbf{E}_{\tau,n}^c)^T) \otimes \mathbf{I}_p \right] \text{vec}(\mathbf{U}_{s,q}) \\
& + n^{-1} \sum_{\rho=1}^q \sum_{j_\rho=(\rho-1)p+1}^{\rho p} \sqrt{n} \lambda_{\rho,n} \hat{w}_{j_\rho,\rho,n} \sqrt{n} \left[|\text{vec}(\mathbf{U}_{s,q})(j_\rho) + \sqrt{n^{-1}} \tilde{\mathbf{b}}(j_\rho)| - |\text{vec}(\mathbf{U}_{s,q})(j_\rho)| \right].
\end{aligned} \quad (48)$$

Next, it is proved that $J_b(\tilde{\mathbf{b}}) - J_b(\mathbf{0})$ converges in distribution to a cost $G_b(\tilde{\mathbf{b}})$, whose minimum will turn out to be the limiting point at which $\check{\mathbf{b}}_n$ converges in distribution as $n \rightarrow \infty$. It follows from (a1) that $\hat{\Sigma}_{s,n} = n^{-1} \mathbf{S}_n \mathbf{S}_n^T$ converges almost surely (a.s.) to Σ_s as $n \rightarrow \infty$, whereas $\hat{\mathbf{C}}_{\tau,n}$ converges in distribution to $\mathbf{U}_{s,q}^T$ (this follows from the asymptotic normality assumption). Then, Slutsky's theorem, e.g., see [14], implies that the first term in (48) converges in distribution to $\tilde{\mathbf{b}}^T (\mathbf{D}_{s,q} \otimes \mathbf{I}_p) \tilde{\mathbf{b}}$. Recalling that the estimation error $\mathbf{e}_{\tau,n}^c$ is assumed to converge to a zero-mean Gaussian distribution with finite covariance, the third term converges in distribution (and in probability) to 0. Taking into account (a1) and that $\mathbf{S}_n \mathbf{S}_n^T = n \hat{\Sigma}_{s,n} = n \Sigma_s + \mathbf{E}_{s,n}$, where $n^{-1} \mathbf{E}_{s,n}$ corresponds to the covariance estimation error, it follows readily that the second term in (48) is equal to

$$\begin{aligned}
2n^{-0.5} \tilde{\mathbf{b}}^T \left[(\hat{\mathbf{C}}_{\tau,n} \mathbf{S}_n \mathbf{S}_n^T \mathbf{U}_{s,p-q}) \otimes \mathbf{I}_p \right] \text{vec}(\mathbf{U}_{s,p-q}) &= 2n^{-0.5} \tilde{\mathbf{b}}^T \left[(\mathbf{U}_{s,q}^T \mathbf{E}_{s,n} \mathbf{U}_{s,p-q}) \otimes \mathbf{I}_p \right] \text{vec}(\mathbf{U}_{s,p-q}) \\
&+ 2\tilde{\mathbf{b}}^T \left[(\mathbf{E}_{\tau,n}^c \Sigma_s \mathbf{U}_{s,p-q}) \otimes \mathbf{I}_p \right] \text{vec}(\mathbf{U}_{s,p-q}) + 2n^{-1} \tilde{\mathbf{b}}^T \left[(\mathbf{E}_{\tau,n}^c \mathbf{E}_{s,n} \mathbf{U}_{s,p-q}) \otimes \mathbf{I}_p \right] \text{vec}(\mathbf{U}_{s,p-q}).
\end{aligned} \quad (49)$$

Recall that $\mathbf{E}_{\tau,n}$ converges in distribution to a zero-mean Gaussian random variables with finite covariance, whereas $\mathbf{E}_{s,n}$ adheres to a Wishart distribution with scaling matrix Σ_s and n degrees of freedom [19, pg.

47]. Then, it follows readily that the first and third terms in (50) converge to zero in distribution (thus in probability too). Then, we have that the lhs of (49) converges to

$$2\tilde{\mathbf{b}} [(\mathbf{E}_c \boldsymbol{\Sigma}_s \mathbf{U}_{s,p-q}) \otimes \mathbf{I}_p] \text{vec}(\mathbf{U}_{s,p-q}) = 2\tilde{\mathbf{b}}^T \text{vec}(\mathbf{U}_{s,p-q} \mathbf{U}_{s,p-q}^T \boldsymbol{\Sigma}_s \mathbf{E}_c^T) \quad (50)$$

where \mathbf{E}_c denotes the Gaussian random matrix at which $\mathbf{E}_{\tau,n}^c$ converges in distribution as $n \rightarrow \infty$. Similarly, we can show that the fourth summand in (48) converges in distribution, as $n \rightarrow \infty$, to

$$2\tilde{\mathbf{b}}^T [(\mathbf{U}_{s,q}^T \boldsymbol{\Sigma}_s \mathbf{E}_c^T) \otimes \mathbf{I}_p] \text{vec}(\mathbf{U}_{s,q}) = 2\tilde{\mathbf{b}}^T \text{vec}(\mathbf{U}_{s,q} \mathbf{E}_c \boldsymbol{\Sigma}_s \mathbf{U}_{s,q}). \quad (51)$$

The limiting noise terms in (50) and (51) are zero-mean and uncorrelated. Now, we examine the limiting behavior of the double sum in (48). If $\text{vec}(\mathbf{U}_{s,q})(j) \neq 0$ then $\lim_{n \rightarrow \infty} \sqrt{n} [|\text{vec}(\mathbf{U}_{s,q})(j) + \sqrt{n^{-1}} \tilde{\mathbf{b}}(j)| - |\text{vec}(\mathbf{U}_{s,q})(j)|] = \text{sgn}[\text{vec}(\mathbf{U}_{s,q})(j)] \tilde{\mathbf{b}}(j)$. Since $\hat{w}_{j,\rho,n}$ converges in probability to $|\text{vec}(\mathbf{U}_{s,q})(j)|^{-\gamma}$, we can deduce that if $\lambda_{\rho,n}$ is selected as suggested by the first limit in (23), then the corresponding term in the double sum in (48) goes to zero in distribution (and in probability) as $n \rightarrow \infty$.

For the case where $\text{vec}(\mathbf{U}_{s,q})(j) = 0$, it holds that $\sqrt{n} [|\text{vec}(\mathbf{U}_{s,q})(j) + \sqrt{n^{-1}} \tilde{\mathbf{b}}(j)| - |\text{vec}(\mathbf{U}_{s,q})(j)|] = |\tilde{\mathbf{b}}(j)|$, and also $\sqrt{n} \lambda_{\rho,n} \hat{w}_{j,\rho,n} = \sqrt{n} \lambda_{\rho,n} n^{\gamma/2} (\sqrt{n} |\text{vec}(\hat{\mathbf{U}}_{s,q})(j))|^{-\gamma}$. Since $\hat{\mathbf{U}}_{s,q}$ is an asymptotically normal estimator for $\mathbf{U}_{s,q}$ it follows that $(\sqrt{n} |\text{vec}(\hat{\mathbf{U}}_{s,q})(j))|^{-\gamma}$ converges in distribution to a random variable of finite variance as $n \rightarrow \infty$. Given that $\lambda_{\rho,n}$ satisfies the second limit in (23), using the previous two limits and Slutsky's theorem we have that the quantity in (48) converges in distribution to

$$G_b(\tilde{\mathbf{b}}) = \tilde{\mathbf{b}}_{\mathcal{S}_o}^T [\mathbf{D}_{s,q} \otimes \mathbf{I}_p]_{\mathcal{S}_o} \tilde{\mathbf{b}}_{\mathcal{S}_o} - 2\tilde{\mathbf{b}}_{\mathcal{S}_o}^T [\text{vec}(\mathbf{U}_{s,p-q} \mathbf{U}_{s,p-q}^T \boldsymbol{\Sigma}_s \mathbf{E}_c^T - \mathbf{U}_{s,q} \mathbf{E}_c \boldsymbol{\Sigma}_s \mathbf{U}_{s,q})]_{\mathcal{S}_o} \quad (52)$$

if $\tilde{\mathbf{b}}(j) = 0$ for all $j \notin \mathcal{S}_o$; otherwise, the limit is ∞ . The notation $[\mathbf{M}]_{\mathcal{S}_o}$ in (52) denotes the submatrix of \mathbf{M} whose row and column indices are in $\mathcal{S}(\mathbf{U}_{s,q})$. The optimal solution of (52) $\check{\mathbf{b}}$ is given by

$$\check{\mathbf{b}}(\mathcal{S}_o) = ([\mathbf{D}_{s,q} \otimes \mathbf{I}_p]_{\mathcal{S}_o})^{-1} [\text{vec}(\mathbf{U}_{s,p-q} \mathbf{U}_{s,p-q}^T \boldsymbol{\Sigma}_s \mathbf{E}_c^T - \mathbf{U}_{s,q} \mathbf{E}_c \boldsymbol{\Sigma}_s \mathbf{U}_{s,q})]_{\mathcal{S}_o}, \text{ and } \check{\mathbf{b}}(\bar{\mathcal{S}}_o) = \mathbf{0}. \quad (53)$$

Since the cost in (8) is strictly convex wrt \mathbf{B} and $J_b(\tilde{\mathbf{b}}) - J_b(\mathbf{0}) \xrightarrow{d} G_b(\tilde{\mathbf{b}})$ as $n \rightarrow \infty$, one can readily apply the epi-convergence results in [22] to establish that $\check{\mathbf{b}}_n \xrightarrow{d} \check{\mathbf{b}}$ as $n \rightarrow \infty$, while $\check{\mathbf{b}}$ corresponds to a zero-mean Gaussian random vector. This establishes asymptotic normality of $\hat{\mathbf{B}}_{\tau,n}$. An interesting thing to notice is that when setting in (8) $\mu = 0$ and $\{\lambda_{\rho,n} = 0\}_{\rho=1}^q$ (standard PCA approach) it follows that the corresponding cost in (48) converges in distribution to the one in (52) with $\mathcal{S}_o = \{1, \dots, p\}$. This result establishes that the covariance of $\check{\mathbf{b}}$ is equal to $[\boldsymbol{\Sigma}_{E_b}]_{\mathcal{S}_o}$, where $\boldsymbol{\Sigma}_{E_b}$ is the limiting covariance matrix of the estimation error when the standard PCA approach is employed.

Next, we prove that the probability of finding the correct support converges to unity as $n \rightarrow \infty$. Letting $\hat{\mathcal{S}}_{B,n} := \mathcal{S}(\hat{\mathbf{B}}_{\tau,n})$, we have to show that: i) $\Pr(\{i \in \hat{\mathcal{S}}_{B,n}\}) \xrightarrow{n \rightarrow \infty} 1 \forall i \in \mathcal{S}_o$; and ii) $\Pr(\{i \in \hat{\mathcal{S}}_{B,n}\}) \xrightarrow{n \rightarrow \infty} 0 \forall i \notin \mathcal{S}_o$.

$0 \forall i \notin \mathcal{S}_o$. The asymptotic normality of $\hat{\mathbf{B}}_{\tau,n}$ implies that $\hat{\mathbf{B}}_{\tau,n} \xrightarrow{d} \mathbf{U}_{s,q}$. Since $\mathbf{U}_{s,q}$ is a constant matrix, it also holds that $\hat{\mathbf{B}}_{\tau,n} \xrightarrow{p} \mathbf{U}_{s,q}^T$ and the first part of the proof is established. Concerning the second part, differentiate the cost in (47) wrt \mathbf{b} , and apply the first-order optimality conditions to obtain an equality whose lhs and right hs (rhs) are normalized with n . It then holds $\forall j \in \{1, \dots, p\}$ and $\rho \in \{1, \dots, q\}$ for which $(\rho - 1)p + j \in \hat{\mathcal{S}}_{B,n}$ that

$$\frac{2 \left[(\hat{\mathbf{C}}_{\tau,n} \mathbf{S}) \otimes \mathbf{I}_p \right]_{(\rho-1)p+j}^T \left[((\hat{\mathbf{C}}_{\tau,n} \mathbf{S})^T \otimes \mathbf{I}_p) [\text{vec}(\mathbf{U}_{s,q}) - \hat{\mathbf{b}}_{\tau,n}] + \mathbf{z}_{q,n} - \sqrt{n^{-1}} [(\mathbf{E}_{c,n} \mathbf{S})^T \otimes \mathbf{I}_p] \text{vec}(\mathbf{U}_{s,q}) \right]}{\sqrt{n}} + \frac{2\mu(\hat{\mathbf{b}}_{\tau,n}((\rho-1)p+j) - \hat{\mathbf{c}}_{\tau,n}^t((\rho-1)p+j))}{\sqrt{n}} = \text{sgn}[\hat{\mathbf{b}}_{\tau,n}((\rho-1)p+j)] \frac{\lambda_{\rho,n} \hat{w}_{j,\rho,n}}{\sqrt{n}} \quad (54)$$

where $[\mathbf{M}]_j$ denotes the j th row of matrix \mathbf{M} and $\hat{\mathbf{b}}_{\tau,n} = \text{vec}(\hat{\mathbf{B}}_{\tau,n})$. The rhs in (54) can be rewritten as $\frac{\lambda_{\rho,n}}{\sqrt{n}} \frac{n^{\gamma/2}}{|\sqrt{n} \hat{\mathbf{U}}_{s,q}(j,\rho)|^\gamma}$ and goes to ∞ when $\{\lambda_{\rho,n}\}_{\rho=1}^q$ are selected according to (23). The second fraction at the lhs of (54) converges to 0 in probability. Next, we show why the first and third terms in the first fraction converge in distribution to a zero-mean Gaussian variable with finite variance. For the first term this is true because: i) $\hat{\mathbf{C}}_{\tau,n} \xrightarrow{p} \mathbf{U}_{s,q}^T$; ii) $n^{-1} \mathbf{S} \mathbf{S}^T \xrightarrow{a.s.} \mathbf{\Sigma}_s$; and iii) $\sqrt{n} [\text{vec}(\mathbf{U}_{s,q}) - \hat{\mathbf{b}}_{\tau,n}]$ converges in distribution to a zero-mean Gaussian random variable as shown earlier. This is the case for the third term too since: i) $n^{-1} \mathbf{S} \mathbf{S}^T \xrightarrow{a.s.} \mathbf{\Sigma}_s$; ii) $\hat{\mathbf{C}}_{\tau,n} \xrightarrow{p} \mathbf{U}_{s,q}^T$; and iii) $\mathbf{E}_{c,n}$ converges in distribution to a zero-mean Gaussian random matrix. Finally, the second term in the first fraction converges in probability to a constant because $n^{-1} \mathbf{S} \mathbf{S}^T \xrightarrow{a.s.} \mathbf{\Sigma}_s$ and $\hat{\mathbf{C}}_{\tau,n} \xrightarrow{p} \mathbf{U}_{s,q}^T$.

Notice that the event $i = (\rho - 1)p + j \in \hat{\mathcal{S}}_{B,n}$ implies equality (54); hence, $\forall i \notin \mathcal{S}_o$ $\Pr(\{i \in \hat{\mathcal{S}}_{B,n}\}) \leq \Pr(\{\text{eq. (54) is true}\})$. However, as $n \rightarrow \infty$ the probability of (54) being satisfied goes to zero since the lhs converges to a Gaussian variable and the rhs goes to ∞ ; thus, $\forall i \notin \mathcal{S}_o$ it holds that $\lim_{n \rightarrow \infty} \Pr(\{i \in \hat{\mathcal{S}}_{B,n}\}) = 0$. \square

REFERENCES

- [1] "NASA images of Mars and all available satellites." [Online]. Available: <http://photojournal.jpl.nasa.gov/catalog/?IDNumber=PIA07890>
- [2] R. G. Baraniuk, E. Candès, R. Nowak, and M. Vetterli, "Compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 12–13, March 2008.
- [3] D. P. Bertsekas, *Nonlinear Programming*. Second Edition, Athena Scientific, 2003.
- [4] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [5] D. R. Brillinger, *Time Series: Data Analysis and Theory*. Expanded Edition, Holden Day, 1981.
- [6] E. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" 2009. [Online]. Available: <http://www.citebase.org/abstract?id=oai:arXiv.org:0912.3599>

- [7] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, Feb. 2006.
- [8] A. D'Aspremont, F. Bach, and L. E. Ghaoui, "Optimal solutions for sparse principal component analysis," *Journal of Machine Learning Research*, vol. 9, pp. 1269–1294, Jul. 2008.
- [9] A. D'Aspremont, L. E. Ghaoui, M. Jordan, and G. Lanckriet, "A direct formulation for sparse PCA using semidefinite programming," *SIAM Review*, vol. 49, no. 3, pp. 434–448, 2007.
- [10] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.
- [11] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. Academic Press, 1983.
- [12] J. E. Fowler, "Compressive-projection principal component analysis," *IEEE Transactions on Image Processing*, vol. 18, pp. 2230–2242, Oct. 2009.
- [13] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.
- [14] G. Gimmert and D. Stirzaker, *Probability and Random Processes*. Third Edition, Oxford, 2001.
- [15] G. H. Golub and C. F. V. Loan, *Matrix Computations*. John Hopkins University Press, 1996.
- [16] V. K. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sep. 2001.
- [17] T. Hastie, R. Tibshirani, and D. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Second Edition, Springer, 2009.
- [18] I. M. Johnstone and A. Lu, "On consistency and sparsity for principal components analysis in high dimensions," *Journal of the American Statistical Association*, vol. 104, no. 486, pp. 682–693, 2011.
- [19] I. Jolliffe, *Principal Component Analysis*. Second Edition, New York: Springer, 2002.
- [20] I. Jolliffe, N. Tendaifilov, and M. Uddin, "A modified principal component technique based on the Lasso," *Journal of Computational and Graphical Statistics*, vol. 12, no. 3, pp. 531–547, 2003.
- [21] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [22] K. Knight and W. Fu, "Asymptotics for Lasso-type estimators," *Annals of Statistics*, vol. 28, no. 5, pp. 1356–1378, 2000.
- [23] C. Leng and H. Wang, "On general adaptive sparse principal component analysis," *Journal of Computational and Graphical Statistics*, vol. 18, no. 1, pp. 201–215, 2009.
- [24] Z. Lu and Y. Zhang, "An augmented Lagrangian approach for sparse principal component analysis," *Mathematical Programming*, 2011 (to appear). [Online]. Available: <http://www.citebase.org/abstract?id=oai:arXiv.org:0907.2079>
- [25] M. K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, pp. 300–303, Dec. 1999.
- [26] D. D. Muresan and T. W. Parks, "Adaptive principal components and image denoising," in *Proc. of the Intl. Conf. on Image Proc.*, vol. I, Barcelona, Spain, Sep. 2003, pp. 101–104.
- [27] H. Shen and J. Z. Huang, "Sparse principal component analysis via regularized low rank matrix approximation," *Journal of Multivariate Statistics*, vol. 99, no. 6, pp. 1015–1034, 2008.
- [28] E. P. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Lecture Notes in Statistics*. Springer-Verlag, 1999.
- [29] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, no. 1, pp. 267–288, 1996.

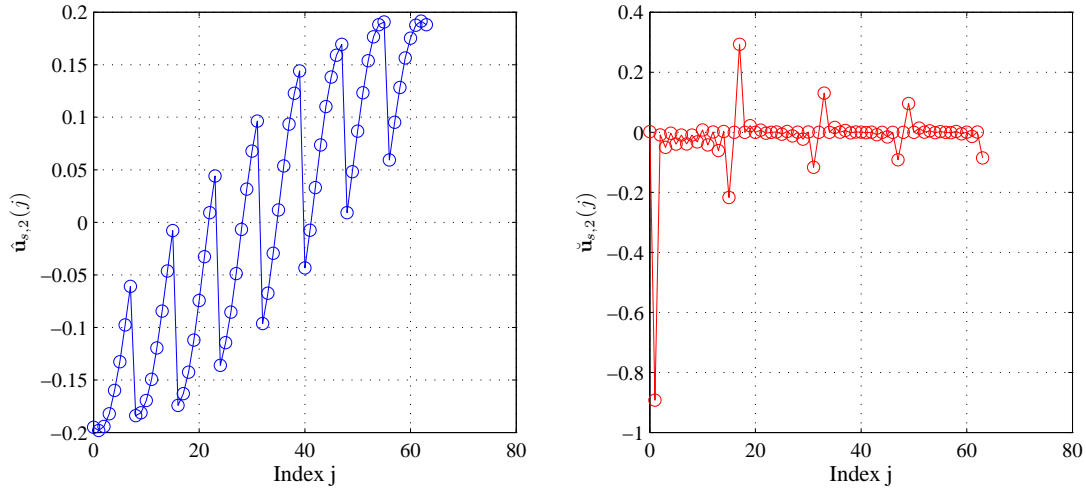


Fig. 1. The second principal eigenvector of $\hat{\Sigma}_s$ (left) and its discrete cosine transform (right).

- [30] P. Tseng, “Convergence of a block coordinate descent method for nondifferentiable minimization,” *Journal of Opt. Theory and Applications*, vol. 109, no. 3, pp. 475–494, Jun. 2001.
- [31] M. O. Ulfarsson and V. Solo, “Sparse variable PCA using geodesic steepest descent,” *IEEE Transactions on Signal Processing*, vol. 10, no. 12, pp. 5823–5832, 2008.
- [32] —, “Sparse variable noisy PCA using ℓ_0 penalty,” in *Proc. of the Intl. Conf. on Acoust., Speech and Sig. Proc.*, Dallas, TX, Mar. 2010, pp. 3950–3953.
- [33] D. M. Witten, R. Tibshirani, and T. Hastie, “A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis,” *Biostatistics*, vol. 10, no. 3, pp. 515–534, 2009.
- [34] B. Yang, “Projection approximation subspace tracking,” *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 95–107, 1995.
- [35] H. Zou, “The adaptive Lasso and its oracle properties,” *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.
- [36] H. Zou, T. Hastie, and R. Tibshirani, “Sparse principal component analysis,” *Journal of Computational and Graphical Statistics*, vol. 15, no. 2, 2006.

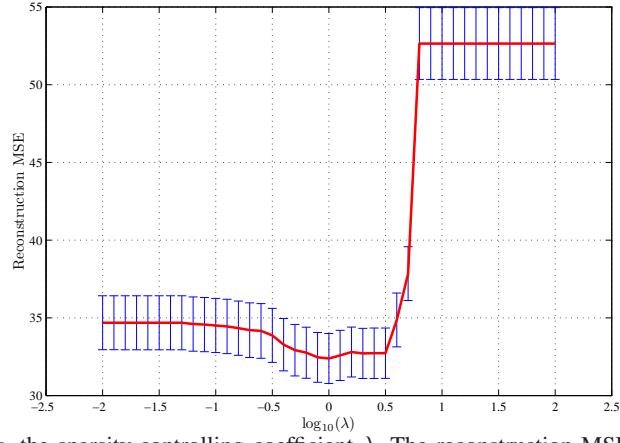


Fig. 2. Reconstruction MSE vs. the sparsity-controlling coefficient λ . The reconstruction MSE is estimated via 5-fold CV.

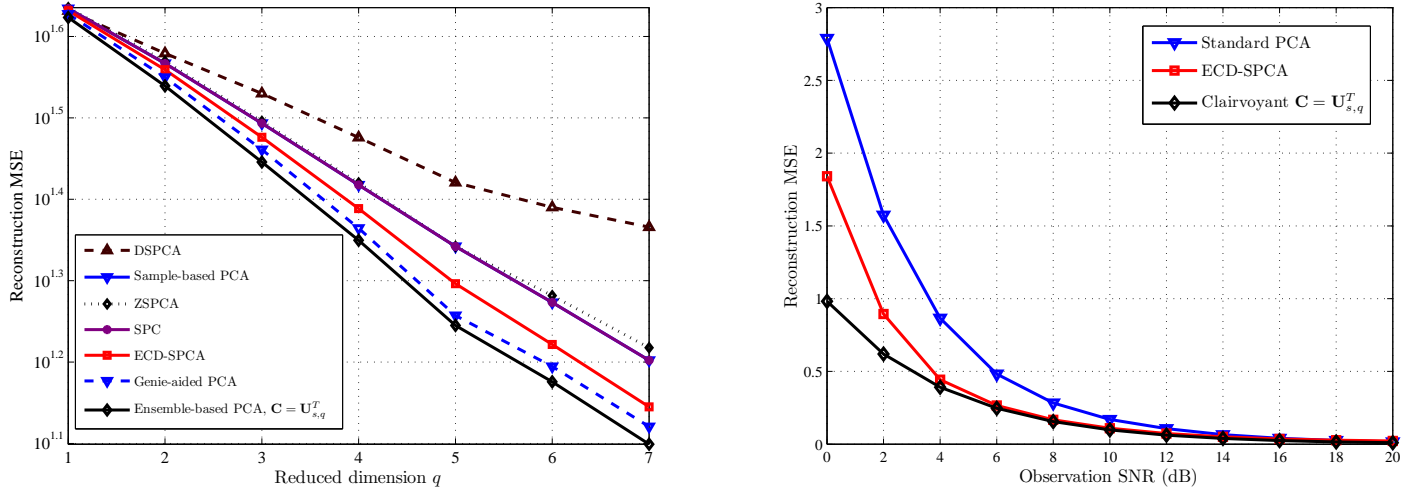


Fig. 3. Reconstruction MSE vs. q in the noiseless case (left); and observation SNR in the noisy scenario with $q = 3$ (right).

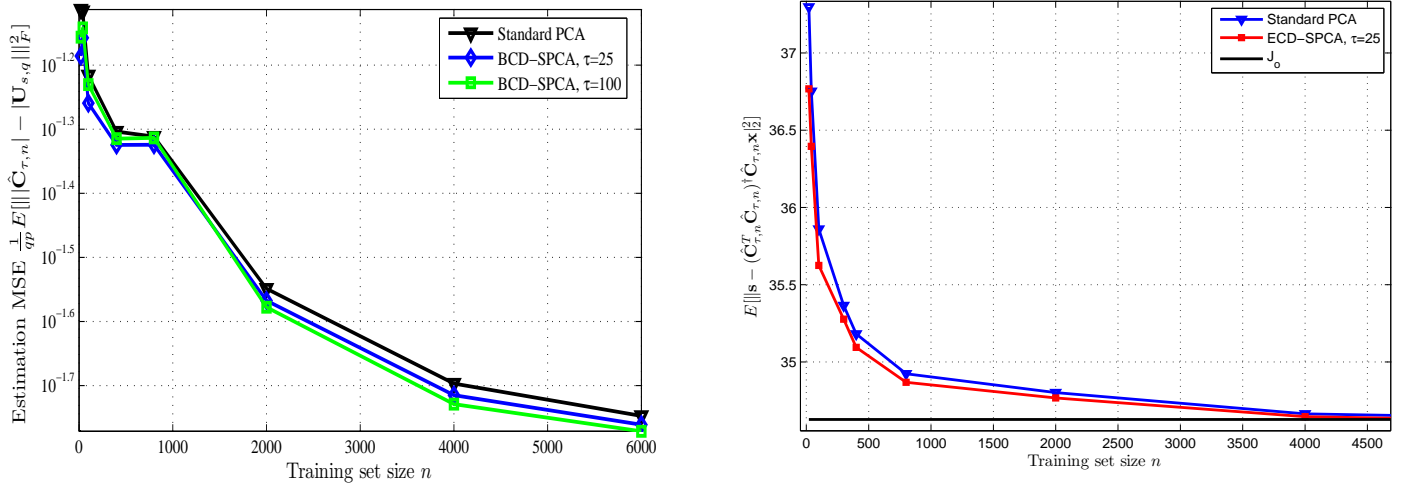


Fig. 4. Normalized estimation MSE $(qp)^{-1} E[\|\hat{\mathbf{C}}_{\tau,n} - \mathbf{U}_{s,q}\|_F^2]$ (left); and reconstruction MSE J_{rec} (right) vs. n for ECD-SPCA with $p = 14$ and $q = 2$.

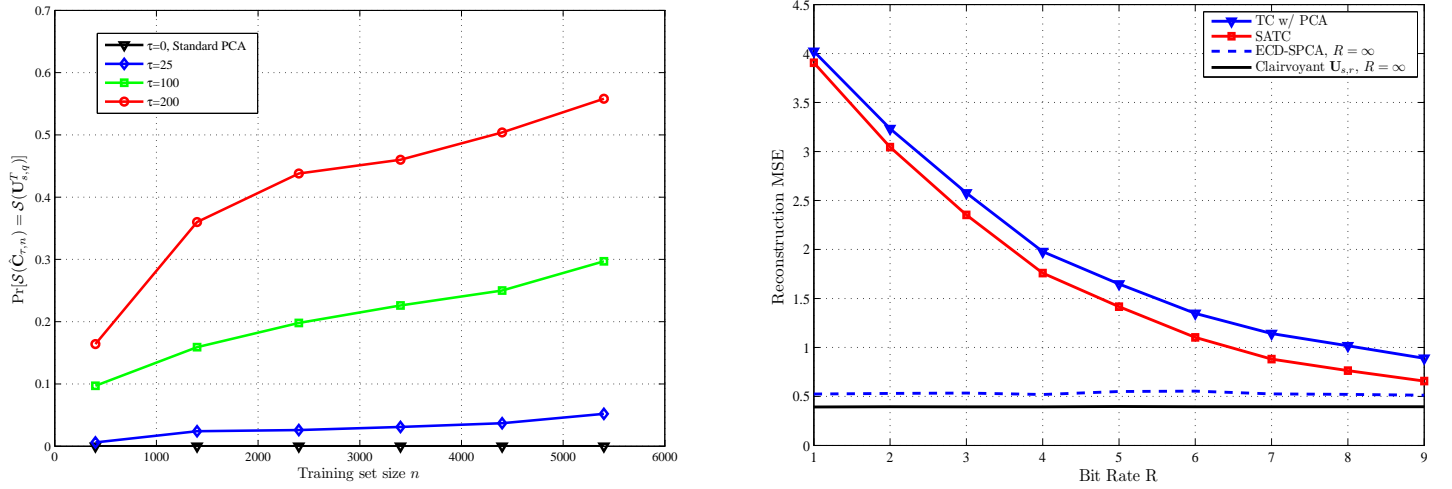


Fig. 5. (Left) Probability that $\hat{\mathbf{C}}_{\tau,n}$ identifies the support of $\mathbf{U}_{s,q}^T$. Variable-size training sets considered in the x-axis, while the different curves correspond to a variable number of coordinate descent recursions; (Right) Reconstruction MSE vs. quantization bit rate R for SATC and a PCA-based TC scheme with $r = q = 3$.

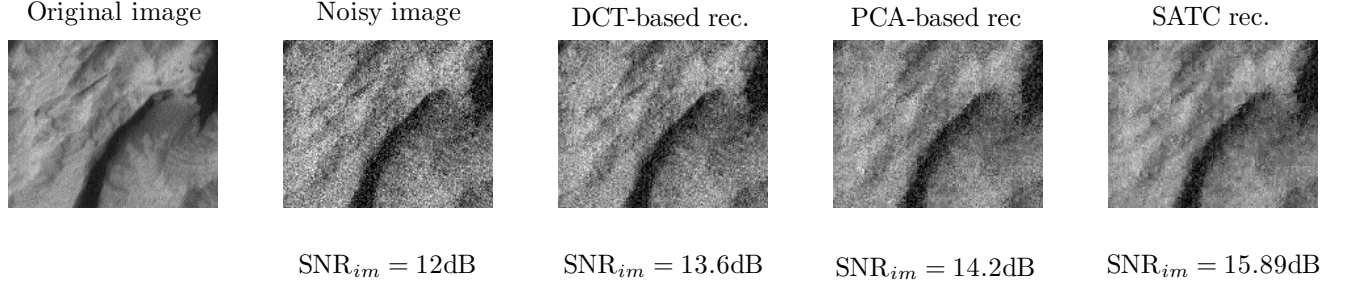


Fig. 6. Original image (leftmost); its noisy version; and its reconstruction using DCT-based TC, PCA-based TC and SATC (right). Reconstructed images are produced after setting $q = 14$ and $R = 7$ bits for each of the 8×8 blocks comprising the original noisy image.

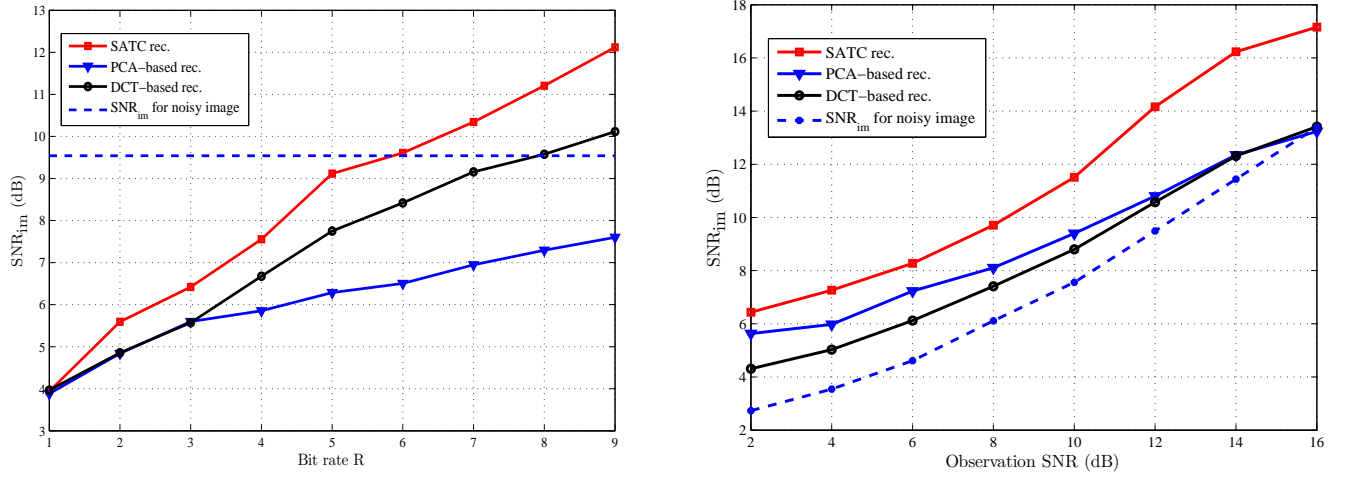


Fig. 7. SNR_{im} vs. bit rate R (left); and observation SNR (right) for different TCs using the images extracted from [1].